

see for hjaace54 - alone

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:29:12 ; Search time 164 Seconds

(without alignments)  
313.653 Million cell updates/sec

Title: HJACE54

Perfect score: 682

Sequence: 1 msrplvpcshalpglspg.....leqlrelriagsvqlcyvhs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

A\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	671	98.4	133	2 AAW61627 Clone HJA
2	671	98.4	133	3 AAY44510 Human Gal
3	671	98.4	133	3 AAB35705 Human gal
4	671	98.4	133	7 AAW01509 Human rec
5	671	98.4	133	7 AAG40247 Human Gal
6	671	98.4	275	3 AAB35706 Human gal
7	671	98.4	275	7 ADG40270 Human Gal
8	671	98.4	296	3 AAB35707 Human gal
9	671	98.4	296	7 ADG40272 Human Gal
10	671	98.4	314	5 AAM48755 Human adi
11	671	98.4	336	3 AAY70029 Human ext
12	124.5	18.3	318	8 ADI13483 Human gal
13	120.5	17.7	171	3 AAY79505 Human car
14	120.5	17.7	196	4 AAB70086 Human sec
15	119	17.4	144	7 ADE57164 Rat Prote
16	119	17.4	144	7 ADE57162 Rat Prote
17	119	17.4	145	7 ADE57162 Rat Prote
18	119	17.4	145	7 ADE57162 Rat Prote
19	118.5	17.4	172	8 ABO04944 Antipsori
20	109.5	16.1	158	6 ABR69609 Human CGD
21	109.5	16.1	258	8 ABO21123 Human car
22	109.5	16.1	258	8 ABR69609 Human car
23	109.5	16.1	258	8 ABR69609 Human car
24	109.5	16.1	316	3 AAY87403 Human PCT
25	109.5	16.1	316	4 AAB85030 Protein e

99 81.5 12.0 323 5 AA015235 Aa015235 Human gal  
100 81.5 12.0 323 5 AAU76211 Aau76211 Human col

## ALIGNMENTS

## RESULT 1

AAW61627

ID AAW61627 standard; protein; 133 AA.

XX AC AAW61627;

XX DT 27-OCT-1998 (first entry)

XX DE Clone HUACE54 of Lectin superfamily.

XX KW Human; receptor; immune disorder; cancers; blood disorder;

XX KW Juvenile rheumatoid arthritis; Graves disease.

XX OS Homo sapiens.

XX PN WO9831799-A2.

XX PD 23-JUL-1998.

XX PF 21-JAN-1998; 98WO-US000959.

XX PR 21-JAN-1997; 97US-0034204P.

XX PR 21-JAN-1997; 97US-0034205P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ni J, Gentz RL, Rosen CA;

XX DR WPI; 1998-427559/36.

XX DR N-PSDB; AAW48123.

XX PT New isolated polynucleotide(s) and encoded receptor polypeptide(s) -

XX PT used to develop products for diagnosing or treating e.g. immune

XX PT disorders, cancers, blood disorders or immuno-compromised disease states.

XX PS Claim 11; Page 53; 79pp; English.

XX CC Clone HUACE54 is a member of the Lectin receptor superfamily. The

XX CC products generated using the receptor can be used for treating abnormal

XX CC conditions related to both an excess of and insufficient amounts of

XX CC receptor activity. They can be used in the treatment of e.g. immune

XX CC disorders, cancers, blood disorders, juvenile rheumatoid arthritis,

XX CC Graves disease or immunocompromised disease states. The products can also

XX CC be used for detection and diagnosis

XX SQ Sequence 133 AA;

Query Match 98.4%; Score 671; DB 2; Length 133;  
Best Local Similarity 99.2%; Pred. No. 6e-76;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKHFTVSLRDOAAHAPVTLRASPADRTL 60

DB 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKHFTVSLRDOAAHAPVTLRASPADRTL 60

QY 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120

DB 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120

QY 121 RISGSVQLYCVHS 133

DB 121 RISGSVQLYCVHS 133

## RESULT 2

AAAY44510

ID AAY44510 standard; protein; 133 AA.

XX AC AAY44510;

XX DT 04-APR-2000 (first entry)

XX DE Human Galectin 11 protein.

XX KW Galectin 11; HUACE54 plasmid; galectin 5; galectin 3; galectin 8;

XX KW cell differentiation; proliferation; metastasis; tumour; apoptosis;

XX KW anti-galectin 11 antibody; treatment; autoimmune disease; cancer;

XX KW inflammatory disease; asthma; allergy; neuropathy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Domain 65..70

XX FT Domain /label= Epitope

XX FT Domain 118..124

XX FT Domain /label= Epitope

XX PN WO200001728-A1.

XX PD 13-JAN-2000.

XX PF 02-JUL-1999; 99WO-US015169.

XX PR 06-JUL-1998; 98US-00109864.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ni J, Gentz RL;

XX DR WPI; 2000-126931/11.

XX DR N-PSDB; AAZ49417.

XX PT New human galectin polypeptide, useful for treating cell growth disorders

XX PT in mammals e.g. cancer, autoimmune diseases, inflammatory diseases,

XX PT asthma and allergic diseases.

XX PS Claim 1; Fig 1; 99pp; English.

XX CC The present sequence is the human galectin 11 protein. It is encoded by

XX CC cDNA derived from HUACE54 plasmid (ATCC 209053). Galectin 11 shows

XX CC homology to rat galectin 5, chicken galectin 3 and human galectin 8. It

XX CC is involved in cell differentiation, proliferation, metastasis of tumour

XX CC cells, cell adhesion, autoimmunity, inflammation, cell-cell substrate

XX CC interactions and apoptosis. The polynucleotides can be used to detect

XX CC cell growth/differentiation disorders and to produce probes and primers

XX CC for detecting polynucleotides and gene mapping. Anti-galectin 11

XX CC antibodies can be used therapeutically and for purification of proteins.

XX CC Galectin 11 can be used to treat cancer, autoimmune diseases,

XX CC inflammatory diseases, asthma, allergy and neuropathies

Query Match 98.4%; Score 671; DB 3; Length 133;  
Best Local Similarity 99.2%; Pred. No. 6e-76;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKHFTVSLRDOAAHAPVTLRASPADRTL 60

DB 1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLPQPKHFTVSLRDOAAHAPVTLRASPADRTL 60

QY 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120

DB 61 AWISRWGQKLLSAPFLFPQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120

QY 121 RISGSVQLYCVHS 133

DB 121 RISGSVQLYCVHS 133

```

RESULT 3
AAB35705
ID AAB35705 standard; protein; 133 AA.
XX
AC AAB35705;
XX
DT 15-FEB-2001 (first entry)
XX
DE Human galectin 11 amino acid sequence.
XX
KW Human; galectin 11; immunosuppressive; thyromimetic; hepatotropic;
KW antiinflammatory; vasotropic; dermatological; nephrotropic; cytostatic;
KW antirheumatic; antiarthritic; antiviral; anti-HIV; neutrotropic; cardiant;
KW neuroprotective; antiParkinsonian; ophthalmological; antianaemic;
KW cerebrotective; antibacterial; immunomodulatory; cancer; allergy;
KW autoimmune disease; inflammatory disease; arthritis; Parkinson's disease;
KW Alzheimer's disease; ischaemia; stroke; anorexia; shock.
XX
OS Homo sapiens.
XX
PN WO200063221-A2.
XX
PD 26-OCT-2000.
XX
PF 21-APR-2000; 2000WO-US010714.
XX
PR 21-APR-1999; 99US-0130390P.
PR 10-DEC-1999; 99US-0169932P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (LJOL-) LAJOLLA INST ALLERGY & IMMUNOLOGY.
XX
PI Ni J, Rosen CA, Gentz RL, Lui F;
XX
DR WPI; 2000-665238/64.
XX
DR N-PSDB; AAC66206.
XX
PT Novel galectin polynucleotides and polypeptides, used in the diagnosis
PT and treatment of cancers, autoimmune disorders, and inflammatory
PT disorders, and to screen for antagonists and agonists.
XX
PS Claim 9; Fig 1; 314pp; English.
XX
CC Polynucleotides AAC66206 - AAC66208 encode human galectin 11, 11alpha and
CC 11beta AAB35705 - AAB35707 respectively. Galectin 11 polynucleotide and
CC protein sequences have immunosuppressive; thyromimetic; hepatotropic;
CC antiinflammatory; vasotropic; dermatological; nephrotropic; antirheumatic
CC ; antiarthritic; antiviral; anti-HIV; neutrotropic; neuroprotective;
CC antiParkinsonian; ophthalmological; cytostatic; antianaemic; cardiant;
CC cerebrotective; antibacterial; and immunomodulatory activity. Galectin
CC 11 polypeptides can be used to treat cell growth disorders, particularly
CC cancer, autoimmune diseases, inflammatory diseases, asthma, or allergic
CC diseases, in a mammal. The proteins and polynucleotides can be
CC administered to regulate cell growth or differentiation in mammals.
CC Antibodies specific for galectin 11 can be used to detect galectin 11
CC polynucleotides and antibodies, used in diagnosis techniques. Proteins,
CC polynucleotides and antibodies, can be used to treat, e.g. Hashimoto's
CC thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease,
CC polymyositis, systemic lupus erythematosus, glomerulonephritis,
CC rheumatoid arthritis, viral infections, inflammation, graft versus host
CC disease, graft rejection, acquired immunodeficiency syndrome, Alzheimer's
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Retinitis
CC pigmentosa, cerebellar degeneration, brain tumour, multiple sclerosis.
CC Sjogren's syndrome, aplastic anaemia, ischaemia, myocardial infarction,
CC stroke, reperfusion injury, liver injury, septic shock, cachexia, and
CC anorexia. PCR primers AAC66211 - AAC66215 are used to amplify galectin 11
CC encoding DNA. Oligonucleotides AAC66216 - AAC66224 are used in the
CC identification and characterisation of the galectin 11 polynucleotides
XX
SQ Sequence 133 AA;

Query Match 98.4%; Score 671; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 6e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPQVIVRGLVLPQPKHFTVSLRQAAHAPVTLRASFDRTL 60
DB 1 MSPRLVPCSHALPQGLSPQVIVRGLVLPQPKHFTVSLRQAAHAPVTLRASFDRTL 60
QY 61 AWISRWGQKLLISAPFLFYPPORFFEVLLIFQEGGLKALNGQGLGATSMNQALEQLREL 120
DB 61 AWISRWGQKLLISAPFLFYPPORFFEVLLIFQEGGLKALNGQGLGATSMNQALEQLREL 120
QY 121 RISGSVQLYCVHS 133
DB 121 RISGSVQLYCVHS 133
RESULT 4
ABW01509
ID ABW01509 standard; protein; 133 AA.
XX
AC ABW01509;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human receptor protein from clone HJACE54.
XX
KW Human; receptor; treatment; inflammation; autoimmunity; gene therapy;
KW tumour; vaccine; immunosuppressive; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003129696-A1.
XX
PD 10-JUL-2003.
XX
PF 29-MAY-2002; 2002US-00156136.
XX
PR 21-JAN-1997; 97US-0034204P.
PR 21-JAN-1997; 97US-0034205P.
PR 21-JAN-1998; 98US-00010146.
PR 19-JAN-2001; 2001US-00764452.
XX
PA (NIJJ/) NI J.
PA (ROSE/) ROSEN C A.
PA (GENTZ/) GENTZ R.
XX
PI Ni J, Rosen CA, Gentz R;
XX
DR WPI; 2003-829564/77.
DR N-PSDB; AAD62619.
XX
PT New polynucleotides and their encoded receptor polypeptides, useful for
PT modulating the activity or expression of receptor polypeptides in a
PT subject, particularly for treating inflammation, autoimmunity or tumors.
XX
PS Claim 10; Page 49; 91pp; English.
XX
CC The present invention provides novel receptor proteins and their encoding
CC nucleotides. The invention is useful for treating a subject in need of
CC enhanced activity or expression of receptor polypeptide. The antagonist/
CC inhibitor of the receptor or polynucleotide encoding the receptor or the
CC polypeptide that competes with the receptor for its ligand is useful for
CC treating a subject having a need to inhibit activity or expression of the
CC receptor polypeptide. The invention is useful in the design of protocols
CC for the treatment of the diseases such as inflammation, autoimmunity and
CC tumours. The invention is also useful in gene therapy and as a vaccine
CC against these diseases. The present sequence is human receptor protein of
CC the present invention
XX
SQ Sequence 133 AA;

Query Match 98.4%; Score 671; DB 7; Length 133;
Best Local Similarity 99.2%; Pred. No. 6e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

```

```
QY      1 MSPRLVPCSHALPQGLSPGVIIVRGLVLPQEPKFTVSLRDOAAHAPVTLRASFAADRTL 60
      |||
Db      1 MSPRLVPCSHALPQGLSPGVIIVRGLVLPQEPKFTVSLRDOAAHAPVTLRASFAADRTL 60
      |||
QY      61 AMISRWGQKKLISAPFLFYPPQRFVFFVLLLFQEGGLKIALNGQGLGATSMNQALEQLEL 120
      |||
Db      61 AMISRWGQKKLISAPFLFYPPQRFVFFVLLLFQEGGLKIALNGQGLGATSMNQALEQLEL 120
      |||
QY      121 RISGSVQLYCVHS 133
      |||
Db      121 RISGSVQLYCVHS 133
      |||

RESULT 5
ADG40247
ID      ADG40247 standard; protein; 133 AA.
XX
AC      ADG40247;
XX
DT      26-FEB-2004 (first entry)
XX
DE      Human Galectin 11 protein.
XX
KW      Galectin 11; cytostatic; immunosuppressive; antiinflammatory;
KW      antiasthmatic; antiallergic; Gene therapy; cell growth disorder; cancer;
KW      autoimmune disease; inflammatory disease; asthma; allergic disease;
KW      human.
XX
OS      Homo sapiens.
XX
PN      US2003208044-A1.
XX
PD      06-NOV-2003.
XX
PF      06-JUN-2003; 2003US-00455366.
XX
PR      21-JAN-1997; 97US-0034204P.
PR      21-JAN-1997; 97US-0034205P.
PR      21-JAN-1998; 98US-00010146.
PR      06-JUL-1998; 98US-00109864.
PR      21-APR-1999; 99US-0130390P.
PR      10-DEC-1999; 99US-0169932P.
PR      21-APR-2000; 2000US-0055170.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
PA      (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.
XX
PI      Ni J, Gentz RL, Rosen CA, Liu F;
XX
DR      WPI: 2003-864797/80.
DR      N-PSDB; ADG40246.
XX
PT      New galectin 11 polypeptides, useful for treating a cell growth disorder,
PT      e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or
PT      allergic diseases.
XX
PS      Claim 9; SEQ ID NO 2; 123pp; English.
XX
CC      The invention relates to an isolated polynucleotide comprising a sequence
CC      that is at least 95% identical to a sequence encoding Human Galectin 11,
CC      galectin 11alpha, galectin 11beta (or their defined fragments). Also
CC      included are a method of making a recombinant vector by inserting the
CC      isolated polynucleotide into a vector, a recombinant vector comprising
CC      the polynucleotide, a genetically engineered host cell comprising the
CC      polynucleotide, a method of producing a galectin 11 polypeptide (by
CC      culturing the genetically engineered host cell under conditions suitable
CC      to produce the polypeptide, and recovering the polypeptide), an isolated
CC      galectin 11 polypeptide comprising any of the amino acid sequences cited
CC      above, a pharmaceutical composition comprising the polypeptide and a
CC      carrier, an isolated antibody that binds specifically to the polypeptide,
CC      a method of detecting a galectin 11 polypeptide in a sample (by
CC      contacting the sample with the antibody, and detecting the presence of
```

```
CC      the antibody bound to the polypeptide), a method of treating a cell
CC      growth disorder in a mammal by administering the polypeptide to the
CC      mammal, and a method of regulating cell growth or differentiation in a
CC      mammal by administering a galectin 11 polypeptide or polynucleotide to
CC      the mammal to suppress cell growth or differentiation. The
CC      polynucleotide, polypeptide and composition are useful for treating a
CC      cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory
CC      diseases, asthma, or allergic diseases (many examples cited in the
CC      specification). The present sequence represents a galectin 11 protein.
XX
SQ      Sequence 133 AA;

      Query Match      98.4%; Score 671; DB 7; Length 133;
      Best Local Similarity 99.2%; Pred. No. 6e-76;
      Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MSPRLVPCSHALPQGLSPGVIIVRGLVLPQEPKFTVSLRDOAAHAPVTLRASFAADRTL 60
      |||
Db      1 MSPRLVPCSHALPQGLSPGVIIVRGLVLPQEPKFTVSLRDOAAHAPVTLRASFAADRTL 60
      |||
QY      61 AMISRWGQKKLISAPFLFYPPQRFVFFVLLLFQEGGLKIALNGQGLGATSMNQALEQLEL 120
      |||
Db      61 AMISRWGQKKLISAPFLFYPPQRFVFFVLLLFQEGGLKIALNGQGLGATSMNQALEQLEL 120
      |||
QY      121 RISGSVQLYCVHS 133
      |||
Db      121 RISGSVQLYCVHS 133
      |||

RESULT 6
AAB35706
ID      AAB35706 standard; protein; 275 AA.
XX
AC      AAB35706;
XX
DT      15-FEB-2001 (first entry)
XX
DE      Human galectin 11alpha amino acid sequence.
XX
KW      Human; galectin 11; immunosuppressive; thyromimetic; hepatotropic;
KW      antiinflammatory; vasotropic; dermatological; nephrotropic; cytostatic;
KW      antirheumatic; antiarthritic; antiviral; anti-HIV; nootropic; cardiant;
KW      neuroprotective; antiParkinsonian; ophthalmological; antianaemic;
KW      cerebroprotective; antibacterial; immunomodulatory; cancer; allergy;
KW      autoimmune disease; inflammatory disease; arthritis; Parkinson's disease;
KW      Alzheimer's disease; ischaemia; stroke; anorexia; shock.
XX
OS      Homo sapiens.
XX
PN      WO200063221-A2.
XX
PD      26-OCT-2000.
XX
PF      21-APR-2000; 2000WO-US010714.
XX
PR      21-APR-1999; 99US-0130390P.
PR      10-DEC-1999; 99US-0169932P.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
PA      (LJOL-) LAJOLLA INST ALLERGY & IMMUNOLOGY.
XX
PI      Ni J, Rosen CA, Gentz RL, Lui F;
XX
DR      WPI: 2000-665238/64.
DR      N-PSDB; AAC66207.
XX
PT      Novel galectin polynucleotides and polypeptides, used in the diagnosis
PT      and treatment of cancers, autoimmune disorders, and inflammatory
PT      disorders, and to screen for antagonists and agonists.
XX
PS      Claim 9; Fig 6; 314pp; English.
XX
CC      Polynucleotides AAC66206 - AAC66208 encode human galectin 11, 11alpha and
```



```

CC 1libeta AAB35705 - AAB35707 respectively. Galectin 11 polynucleotide and
CC protein sequences have immunosuppressive; thyromimetic; hepatotropic;
CC antiinflammatory; vasotropic; dermatological; nephrotropic; antirheumatic
CC ; antiarthritic; antiviral; anti-HIV; nootropic; neuroprotective;
CC antiParkinsonian; ophthalmological; cytostatic; antiansemetic; cardiac;
CC cerebroprotective; antibacterial; and immunomodulatory activity. Galectin
CC 11 polypeptides can be used to treat cell growth disorders, particularly
CC cancer, autoimmune diseases, inflammatory diseases, asthma, or allergic
CC diseases, in a mammal. The proteins and polynucleotides can be
CC administered to regulate cell growth or differentiation in mammals.
CC Antibodies specific for galectin 11 can be used to detect galectin 11
CC polypeptides in a samples, used in diagnosis techniques. Proteins,
CC polynucleotides and antibodies, can be used to treat, e.g. Hashimoto's
CC thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease,
CC polymyositis, systemic lupus erythematosus, glomerulonephritis,
CC rheumatoid arthritis, viral infections, inflammation, graft versus host
CC disease, graft rejection, acquired immunodeficiency syndrome, Alzheimer's
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Retinitis
CC pigmentosa, cerebellar degeneration, brain tumour, multiple sclerosis,
CC Sjogren's syndrome, aplastic anaemia, ischaemia, myocardial infarction,
CC stroke, reperfusion injury, liver injury, septic shock, cachexia, and
CC anorexia. PCR primers AAC66211 - AAC66215 are used to amplify galectin 11
CC encoding DNA. Oligonucleotides AAC66216 - AAC66224 are used in the
CC identification and characterisation of the galectin 11 polynucleotides
XX
SQ Sequence 275 AA;

Query Match          98.4%; Score 671; DB 3; Length 275;
Best Local Similarity 99.2%; Pred. No. 1.6e-75;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPQGVIIVRGLVLQEPKHTVSLRDQAAHAPVTLRASFDRTL 60
DB 143 MSPRLEVPCHALPQGLSPQGVIIVRGLVLQEPKHTVSLRDQAAHAPVTLRASFDRTL 202

QY 61 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120
DB 203 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 262

QY 121 RIGSVQLYCVHS 133
DB 263 RIGSVQLYCVHS 275

RESULT 7
ADG40270
ID ADG40270 standard; protein; 275 AA.
XX
AC ADG40270;
XX
XX 26-FEB-2004 (first entry)
XX
DE Human Galectin 11alpha protein.
XX
XX Galectin 11; cytostatic; immunosuppressive; antiinflammatory;
XX antiasthmatic; antiallergic; gene therapy; cell growth disorder; cancer;
XX autoimmune disease; inflammatory disease; asthma; allergic disease;
XX human.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 1. 121
XX Region /note= "Claimed in claim 9"
XX FT 151. .275
XX FT Region /note= "Claimed in claim 9"
XX
XX US2003208044-A1.
XX
XX 06-NOV-2003.
XX
XX 06-JUN-2003; 2003US-00455366.
XX
XX
XX

```

```

PR 21-JAN-1997; 97US-0034204P.
PR 21-JAN-1997; 97US-0034205P.
PR 21-JAN-1998; 98US-00010146.
PR 06-JUL-1998; 98US-00109864.
PR 21-APR-1999; 99US-0130390P.
PR 10-DEC-1999; 99US-0169932P.
PR 21-APR-2000; 2000US-00557170.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.
PA
XX Ni J, Gentz RL, Rosen CA, Liu F;
XX WPI; 2003-864797/80.
XX N-PSDB; ADG40269.
DR
XX New galectin 11 polypeptides, useful for treating a cell growth disorder,
XX e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or
XX allergic diseases.
XX
XX Claim 9; SEQ ID NO 25; 123pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising a sequence
XX that is at least 95% identical to a sequence encoding Human Galectin 11,
XX galectin 11alpha, galectin 11beta (or their defined fragments). Also
XX included are a method of making a recombinant vector by inserting the
XX isolated polynucleotide into a vector, a recombinant vector comprising the
XX polynucleotide, a genetically engineered host cell comprising the
XX polynucleotide, a method of producing a galectin 11 polypeptide (by
XX culturing the genetically engineered host cell under conditions suitable
XX to produce the polypeptide, and recovering the polypeptide), an isolated
XX galectin 11 polypeptide comprising any of the amino acid sequences cited
XX above, a pharmaceutical composition comprising the polypeptide and a
XX carrier, an isolated antibody that binds specifically to the polypeptide,
XX a method of detecting a galectin 11 polypeptide in a sample (by
XX contacting the sample with the antibody, and detecting the presence of
XX the antibody bound to the polypeptide), a method of treating a cell
XX growth disorder in a mammal by administering the polypeptide to the
XX mammal, and a method of regulating cell growth or differentiation in a
XX mammal by administering a galectin 11 polypeptide or polynucleotide to
XX the mammal to suppress cell growth or differentiation. The
XX polynucleotide, polypeptide and composition are useful for treating a
XX cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory
XX diseases, asthma, or allergic diseases (many examples cited in the
XX specification). The present sequence represents a galectin 11 protein.
XX
SQ Sequence 275 AA;

Query Match          98.4%; Score 671; DB 7; Length 275;
Best Local Similarity 99.2%; Pred. No. 1.6e-75;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLEVPCHALPQGLSPQGVIIVRGLVLQEPKHTVSLRDQAAHAPVTLRASFDRTL 60
DB 143 MSPRLEVPCHALPQGLSPQGVIIVRGLVLQEPKHTVSLRDQAAHAPVTLRASFDRTL 202

QY 61 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 120
DB 203 AWISRWQKKLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQALEQREL 262

QY 121 RIGSVQLYCVHS 133
DB 263 RIGSVQLYCVHS 275

RESULT 8
AAB35707
ID AAB35707 standard; protein; 296 AA.
XX
XX AAB35707;
XX
XX 15-FEB-2001 (first entry)
XX
XX

```

DE Human galectin 11beta amino acid sequence.

XX Human; galectin 11; immunosuppressive; thyromimetic; hepatotropic;  
KW antiinflammatory; vasotropic; dermatological; nephrotropic; cytostatic;  
KW antirheumatic; antiarthritic; antiviral; anti-HIV; nootropic; cardiant;  
KW neuroprotective; antiParkinsonian; ophthalmological; antianaemic;  
KW cerebroprotective; antibacterial; immunomodulatory; cancer; allergy;  
KW autoimmune disease; inflammatory disease; arthritis; Parkinson's disease;  
KW Alzheimer's disease; ischaemia; stroke; anorexia; shock.

XX Homo sapiens.

XX WO200063221-A2.

XX 26-OCT-2000.

XX 21-APR-2000; 2000WO-US010714.

PR 21-APR-1999; 99US-0130390P.

PR 10-DEC-1999; 99US-0169932P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (LJOL-) LAJOLLA INST ALLERGY & IMMUNOLOGY.

XX Ni J, Rosen CA, Gentz RL, Lui F;

XX WPI; 2000-665238/64.

DR N-PSDB; AAC66208.

XX Novel galectin polynucleotides and polypeptides, used in the diagnosis  
PT and treatment of cancers, autoimmune disorders, and inflammatory  
PT disorders, and to screen for antagonists and agonists.

XX Claim 9; Fig 7; 314pp; English.

XX Polynucleotides AAC66206 - AAC66208 encode human galectin 11, 11alpha and  
CC 11beta AAB35705 - AAB35707 respectively. Galectin 11 polynucleotide and  
CC protein sequences have immunosuppressive; thyromimetic; hepatotropic;  
CC antiinflammatory; vasotropic; dermatological; nephrotropic; antirheumatic  
CC ; antiarthritic; antiviral; anti-HIV; nootropic; neuroprotective;  
CC antiParkinsonian; ophthalmological; cytostatic; antianaemic; cardiant;  
CC cerebroprotective; antibacterial; and immunomodulatory activity. Galectin  
CC 11 polypeptides can be used to treat cell growth disorders, particularly  
CC cancer, autoimmune diseases, inflammatory diseases, asthma, or allergic  
CC diseases, in a mammal. The proteins and polynucleotides can be  
CC administered to regulate cell growth or differentiation in mammals.  
CC Antibodies specific for galectin 11 can be used to detect galectin 11  
CC polypeptides in a samples, used in diagnosis techniques. Proteins,  
CC polynucleotides and antibodies, can be used to treat, e.g. Hashimoto's  
CC thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease,  
CC polymyositis, systemic lupus erythematosus, glomerulonephritis,  
CC rheumatoid arthritis, viral infections, inflammation, graft versus host  
CC disease, graft rejection, acquired immunodeficiency syndrome, Alzheimer's  
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Retinitis  
CC pigmentosa, cerebellar degeneration, brain tumour, multiple sclerosis,  
CC Sjogren's syndrome, aplastic anaemia, ischaemia, myocardial infarction,  
CC stroke, reperfusion injury, liver injury, septic shock, cachexia, and  
CC anorexia. PCR primers AAC66211 - AAC66215 are used to amplify galectin 11  
CC encoding DNA. Oligonucleotides AAC66216 - AAC66224 are used in the  
CC identification and characterisation of the galectin 11 polynucleotides

XX Sequence 296 AA;

Query Match 98.4%; Score 671; DB 3; Length 296;  
Best Local Similarity 99.2%; Pred. No. 1.8e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLRVPCHALPQGLSPGVIIIVGLVLOEPKHTVSRDQAAPVTLRASFDRTL 60  
|||||

Db 164 MSRLRVPCHALPQGLSPGVIIIVGLVLOEPKHTVSRDQAAPVTLRASFDRTL 223  
|||||

QY 61 AMISRWGQKKLISAPFLFYPPQRFVLLLFQEGGLKALNGQGLGATSMNQALEQLREL 120  
|||||

Db 224 AMISRWGQKKLISAPFLFYPPQRFVLLLFQEGGLKALNGQGLGATSMNQALEQLREL 283

QY 121 RIGGSVOLYCVHS 133  
|||||

Db 284 RIGGSVOLYCVHS 296  
|||||

RESULT 9  
ADG40272  
ID ADG40272 standard; protein; 296 AA.  
XX  
AC ADG40272;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Human Galectin 11beta protein.  
XX  
KW Galectin 11; cytostatic; immunosuppressive; antiinflammatory;  
KW antiasthmatic; antiallergic; gene therapy; cell growth disorder; cancer;  
KW autoimmune disease; inflammatory disease; asthma; allergic disease;  
KW human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 1..142  
FT /note= "Claimed in claim 9"  
XX  
PN US2003208044-A1.  
XX  
PD 06-NOV-2003.  
XX  
PF 06-JUN-2003; 2003US-00455366.  
XX  
PR 21-JAN-1997; 97US-0034204P.  
PR 21-JAN-1997; 97US-0034205P.  
PR 21-JAN-1998; 98US-00010146.  
PR 06-JUL-1998; 98US-00109864.  
PR 21-APR-1999; 99US-0130390P.  
PR 10-DEC-1999; 99US-0169932P.  
PR 21-APR-2000; 2000US-00557170.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.  
XX  
PI Ni J, Gentz RL, Rosen CA, Lui F;  
DR WPI; 2003-864797/80.  
DR N-PSDB; ADG40271.  
XX  
PT New galectin 11 polypeptides, useful for treating a cell growth disorder,  
PT e.g. cancer, autoimmune diseases, inflammatory diseases, asthma, or  
PT allergic diseases.

PS Claim 9; SEQ ID NO 27; 123pp; English.

CC The invention relates to an isolated polynucleotide comprising a sequence  
CC that is at least 95% identical to a sequence encoding Human Galectin 11,  
CC galectin 11alpha, galectin 11beta (or their defined fragments). Also  
CC included are a method of making a recombinant vector by inserting the  
CC isolated polynucleotide into a vector, a recombinant vector comprising  
CC the polynucleotide, a genetically engineered host cell comprising the  
CC polynucleotide, a method of producing a galectin 11 polypeptide (by  
CC culturing the genetically engineered host cell under conditions suitable  
CC to produce the polypeptide, and recovering the polypeptide), an isolated  
CC galectin 11 polypeptide comprising any of the amino acid sequences cited  
CC above, a pharmaceutical composition comprising the polypeptide and a  
CC carrier, an isolated antibody that binds specifically to the polypeptide,  
CC a method of detecting a galectin 11 polypeptide in a sample (by  
CC contacting the sample with the antibody, and detecting the presence of  
CC the antibody bound to the polypeptide), a method of treating a cell  
CC growth disorder in a mammal by administering the polypeptide to the  
CC mammal, and a method of regulating cell growth or differentiation in a

CC mammal by administering a galectin 11 polypeptide or polynucleotide to  
CC the mammal to suppress cell growth or differentiation. The  
CC polynucleotide, polypeptide and composition are useful for treating a  
CC cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory  
CC diseases, asthma, or allergic diseases (many examples cited in the  
CC specification). The present sequence represents a galectin 11 protein.  
SQ Sequence 296 AA;

Query Match 98.4%; Score 671; DB 7; Length 296;  
Best Local Similarity 99.2%; Pred. No. 1.9e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVQEPKHTVSLRDQAAHAPVTLRASFADRTL 60  
Db 164 MSPRLVPCSHALPQGLSPGQVIIVRGLVQEPKHTVSLRDQAAHAPVTLRASFADRTL 223

Qy 61 AWISRWGOKKLISAPFLFYQRFVLLFOEGGLKALNGQGLGATSMNQALEQLREL 120  
Db 224 AWISRWGOKKLISAPFLFYQRFVLLFOEGGLKALNGQGLGATSMNQALEQLREL 283

Qy 121 RISGSVOLYCVHS 133  
Db 284 RISGSVOLYCVHS 296

RESULT 10  
AAM48755  
ID AAM48755 standard; protein; 314 AA.  
XX AAM48755;  
AC AAM48755;  
DT 03-APR-2002 (first entry)  
XX Human adipose tissue specific polypeptide SEQ ID NO 2.  
DE Human adipose tissue  
XX Human; adipose tissue.  
KW Homo sapiens.  
OS Homo sapiens.  
XX JP2001309787-A.  
PN JP2001309787-A.  
XX 06-NOV-2001.  
PD (SUMO) SUMITOMO CHEM CO LTD.  
XX WPI; 2002-134851/18.  
DR N-PSDB; ABA96479.  
XX Identification of adipose tissue comprises detecting, in a test tissue  
PT sample, an expression product of a sequence found in adipose tissue.  
XX Claim 1; Page 12-13; 17pp; Japanese.  
PS  
XX The invention relates to identification of adipose tissue comprising that  
CC an expression product of a gene having a sequence encoding a sequence of  
CC 31 amino acids (AAM48755), given in the specification, is detected. The  
CC method is used for the detection of adipose tissue. The present sequence  
CC is that of an adipose specific sequence of the invention  
XX  
SQ Sequence 314 AA;

Query Match 98.4%; Score 671; DB 5; Length 314;  
Best Local Similarity 99.2%; Pred. No. 1.9e-75;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVQEPKHTVSLRDQAAHAPVTLRASFADRTL 60  
Db 182 MSPRLVPCSHALPQGLSPGQVIIVRGLVQEPKHTVSLRDQAAHAPVTLRASFADRTL 241

Qy 61 AWISRWGOKKLISAPFLFYQRFVLLFOEGGLKALNGQGLGATSMNQALEQLREL 120  
Db 242 AWISRWGOKKLISAPFLFYQRFVLLFOEGGLKALNGQGLGATSMNQALEQLREL 301

Qy 121 RISGSVOLYCVHS 133  
Db 302 RISGSVOLYCVHS 314

RESULT 11  
AAV70029  
ID AAV70029 standard; protein; 336 AA.  
XX AAV70029;  
AC AAV70029;  
DT 05-JUN-2000 (first entry)  
XX Human extracellular adhesive protein, EXADH1.  
DE Human; extracellular adhesive protein; EXADH1; cytostatic; antiarthritic;  
KW antiasthmatic; immunosuppressive; antiarteriosclerotic; diagnosis;  
KW treatment; prevention; cancer; leukaemia; melanoma; immune disorder;  
KW rheumatoid arthritis; asthma; atherosclerosis; developmental disorder;  
KW anaemia; epilepsy; Cushing's syndrome; cell growth disorder;  
KW cell differentiation; embryogenesis; morphogenesis; drug screening.  
XX Homo sapiens.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
PH Modified-site 24  
FT /note= "Casein kinase II phosphorylation site"  
FT Region 48..174  
FT /note= "contains at least 46 out of 55 conserved CRD  
FT residues"  
FT Domain 67..174  
FT /note= "this region is similar to galectin carbohydrate  
FT recognition domain (CRD)"  
FT Modified-site 102  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 132  
FT /note= "cAMP and cGMP-dependent protein kinase  
FT phosphorylation site"  
FT Modified-site 164  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 205  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 242  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 242  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 253  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 257  
FT /note= "Casein kinase II phosphorylation site"  
XX WO200009690-A1.  
XX 24-FEB-2000.  
XX 09-AUG-1999; 99WO-US017997.  
XX 10-AUG-1998; 98US-00131648.  
XX (INCY-) INCYTE PHARM INC.  
XX Hillman JJ, Yue H, Corley NC, Guegler KJ, Patterson C;  
XX WPI; 2000-224335/19.  
DR N-PSDB; AAZ50940.  
XX New human extracellular adhesive polypeptide and polynucleotide useful  
PT for diagnosis, prevention and treatment of cancer, immune disorders and



[illegible]





CC growth disorder in a mammal by administering the polypeptide to the  
CC mammal, and a method of regulating cell growth or differentiation in a  
CC mammal by administering a galectin 11 polypeptide or polynucleotide to  
CC the mammal to suppress cell growth or differentiation. The  
CC polynucleotide, polypeptide and composition are useful for treating a  
CC cell growth disorder, e.g. cancer, autoimmune diseases, inflammatory  
CC diseases, asthma, or allergic diseases (many examples cited in the  
CC specification). The present sequence represents a homologue of a galectin  
CC 11 protein.  
XX  
SQ Sequence 145 AA;

Query Match 17.4%; Score 119; DB 7; Length 145;  
Best Local Similarity 26.9%; Pred. No. 2.4e-06;  
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;  
QY 3 PRLEVPFCSHALPQGLSPGQVIVRGVLQPKHPTVSLR---DQAAHAPVTLRASFADRT 59  
DB 11 PNLAVPFTTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70  
QY 60 LAWISRWG-QKKLISAPFLFYPPQRFVLLFOEGGKLKALNGQGLGATSMNQALEQLR 118  
DB 71 TQINNSWGPERSLPGSMPSRQRFVSWILCEGHCFKVAVDGQHCEYSHRLMNLDPIN 130  
QY 119 ELRISGSVOL 128  
DB 131 TLEVAGDIQL 140

RESULT 19  
ADN04944  
ID ADN04944 standard; protein; 172 AA.  
XX  
AC ADN04944;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Antipsoriatic protein sequence #652.  
XX  
KW antipsoriatic; gene therapy; psoriasis; diagnosis.  
XX  
OS Homo sapiens.  
XX  
PN WO2004028479-A2.  
XX  
PD 08-APR-2004.  
XX  
PF 25-SEP-2003; 2003WO-US030907.  
XX  
PR 25-SEP-2002; 2002US-0414006P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;  
PI Wu TD;  
XX  
DR WPI; 2004-305105/28.  
DR N-PSDB; ADN04943.  
XX  
PT New PRO nucleic acid or polypeptide, useful for preparing a  
PT pharmaceutical composition for diagnosing or treating psoriasis in a  
PT mammal.  
XX  
PS Claim 9; SEQ ID NO 1338; 3069pp; English.  
XX  
CC The invention relates to novel polynucleotide and polypeptides for  
CC treating psoriasis or a sequence having at least 80% identity to the  
CC above sequences. The nucleic acid is useful for preparing a composition  
CC for diagnosing or treating psoriasis in a mammal. This sequence  
CC corresponds to one of the polypeptides of the invention.  
XX  
SQ Sequence 172 AA;

Query Match 17.4%; Score 118.5; DB 8; Length 172;  
Best Local Similarity 29.4%; Pred. No. 3.6e-06;  
Matches 40; Conservative 27; Mismatches 58; Indels 11; Gaps 5;  
QY 3 PRLEVP-CSHALPQGLSPGQVIVRGVLQPKHPTVSL-----RDQAAHAPVTLRASF 56  
DB 34 PRLIVPFCGH-IKGGMWPGKVKVIMGIVDLNPESFAISLTGDSDEPPADVAIELKAVFT 92  
QY 57 DRTL---AMIS-RWGQKLIAPFLFYPPQRFVLLFOEGGKLKALNGQGLGATSMNQ 112  
DB 93 DRQLLNSCISGERGEEQAIPYFPIDQPRFVEILCEYPRFRVFDGHLQJDFYHRIQ 152  
QY 113 ALEQLRELIRISGVOL 128  
DB 153 TLSAIDTIKINGDLQI 168  
RESULT 20  
ABR69609  
ID ABR69609 standard; protein; 158 AA.  
XX  
AC ABR69609;  
XX  
DT 12-AUG-2003 (first entry)  
XX  
DE Human CGDD-9 protein.  
XX  
KW Human; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic;  
KW antiinflammatory; gynaecological; cancer; atherosclerosis; epilepsy;  
KW Huntington's disease; stroke; AIDS; allergy; placenta; reproductive;  
KW CGDD; cell growth; cell differentiation; cell death.  
XX  
OS Homo sapiens.  
XX  
PN WO2003027263-A2.  
XX  
PD 03-APR-2003.  
XX  
PF 26-SEP-2002; 2002WO-US031095.  
XX  
PR 28-SEP-2001; 2001US-0326389P.  
PR 05-OCT-2001; 2001US-0327380P.  
PR 05-OCT-2001; 2001US-0328186P.  
PR 12-OCT-2001; 2001US-0329690P.  
PR 26-OCT-2001; 2001US-0345384P.  
PR 26-OCT-2001; 2001US-0348165P.  
PR 02-NOV-2001; 2001US-0350219P.  
PR 09-NOV-2001; 2001US-0344518P.  
PR 09-NOV-2001; 2001US-0345143P.  
PR 16-NOV-2001; 2001US-0332375P.  
PR 03-DEC-2001; 2001US-0336908P.  
PR 07-DEC-2001; 2001US-0340747P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Chawla NK;  
PI Elliott JS, Emerling BM, Gandhi AR, Gietzen KJ, Gorvad AE;  
PI Griffin VA, Hafalia AJA, Ison CH, Kable AE, Kalafus DP;  
PI Lehr-Mason PM, Lu DAM, Marquis JP, Nguyen DB, Ramkumar J;  
PI Richardson TW, Sapperstein SK, Swarnakar A, Tang YT, Tran UK;  
PI Warren BA, Xu Y, Yao MG, Yue H, Yue H;  
XX  
XX WPI; 2003-421159/39.  
DR N-PSDB; ACC90586.  
XX  
CC New human proteins associated with cell growth, differentiation, and  
CC death (CGDD), useful for diagnosing, treating and preventing diseases or  
CC conditions associated with the aberrant CGDD expression e.g. cancer,  
CC AIDS, or epilepsy.  
XX  
PS Claim 1; Page 257; 350pp; English.



CC The invention relates to an isolated polypeptide associated with cell growth, differentiation and death (CGDD). Also disclosed are the polynucleotides encoding the polypeptides. The polypeptides and polynucleotides are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or over expression of CGDD. Such diseases include cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and reproductive disorders, or disorders of the placenta. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of CGDD. The CGDD or its fragments are useful in screening compounds for effectiveness as an agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. Microarrays consisting of polynucleotides of the invention are useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. Sequences given in records AAR69601-AAR69657 represent CGDD polypeptides of the invention

XX SQ Sequence 158 AA;

Query Match 16.1%; Score 109.5; DB 6; Length 158;  
Best Local Similarity 25.5%; Pred. No. 4.3e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLEVPCHALPQGLSPGQVIVRGVLQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 21 TPQLSLPFAARLNTPMGFGTGVVVKGEVNAKSNFVDLLAGSKDIALHLPRLNIKAF 80

QY 57 DRTLAWISRWGQKLLISAPFLFPQRFVLLLFQEGGLKLALNGOGLGATSMNQALEQ 116  
Db 81 VRNSFLOESNGEERNITSPFSPGMYFEMIYCDVREKAVNGVHSLVKHFKELSS 140

QY 117 LRELIRSGSVOLYCVHS 133  
Db 141 IDTLEINGDIHLLEVR 157

RESULT 21  
ID ADO21123  
AC ADO21123;  
XX 12-AUG-2004 (first entry)  
DE Human carbohydrate-associated protein (CHOP) #19.  
KW human; carbohydrate-associated protein; CHOP; neurodegenerative disorder;  
KW Parkinson's disease; Alzheimer's disease; muscular disorder;  
KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;  
KW Grave's disease; cancer; leukemia; immunological disorder; scleroderma;  
KW systemic lupus erythematosus; allergy; gastrointestinal disorder;  
KW Crohn's disease; renal disorder; Goodpasture's syndrome; infection;  
KW cardiovascular disorder; atherosclerosis; hepatic disease; cirrhosis.  
XX  
OS Homo sapiens.  
XX WO200404166-A2.  
XX 27-MAY-2004.  
XX 10-NOV-2003; 2003WO-US035947.  
XX 12-NOV-2002; 2002US-0425423P.  
XX 21-JAN-2003; 2003US-0441847P.  
XX 10-MAR-2003; 2003US-0453882P.  
XX 20-MAR-2003; 2003US-0456645P.  
XX 16-APR-2003; 2003US-0463676P.  
XX (INCY-) INCYTE CORP.  
XX

PI Chawla NK, Tang YT, Griffin JA, Yang YG, Ramkumar J, Khare R;  
PI Richardson TW, Becha SD, Tran UK, Kable AE, Swarnakar A, Warren BA;  
PI Elliott VS, Marquis JP, Hafalia AJA;  
XX  
DR WPI: 2004-420308/39.  
DR N-PSDB; ADO21143.  
XX  
PT New CHOP polypeptides, useful for diagnosing, preventing, and treating disorders associated with abnormal expression or activity of CHOP, e.g. neuromuscular, immunological, cardiovascular disorders, cancer and/or infections.  
XX  
PS Claim 1; SEQ ID NO 19; 181pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of human carbohydrate-associated proteins (CHOP). The DNA and protein sequences of the invention are useful for diagnosing, preventing and treating disorders associated with abnormal expression or activity of CHOP, such as: neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy, catatonias), endocrine disorders (e.g. diabetes, Grave's disease), cancers (e.g. leukemia), immunological disorders (e.g. scleroderma, systemic lupus erythematosus, allergies), gastrointestinal disorders (e.g. Crohn's disease), renal disorders (e.g. Goodpasture's syndrome), infections (e.g. viral, bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis). The present amino acid sequence represents a human CHOP protein of the invention.  
XX  
SQ Sequence 258 AA;

Query Match 16.1%; Score 109.5; DB 8; Length 258;  
Best Local Similarity 25.5%; Pred. No. 8.4e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLEVPCHALPQGLSPGQVIVRGVLQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 121 TPQLSLPFAARLNTPMGFGTGVVVKGEVNAKSNFVDLLAGSKDIALHLPRLNIKAF 180

QY 57 DRTLAWISRWGQKLLISAPFLFPQRFVLLLFQEGGLKLALNGOGLGATSMNQALEQ 116  
Db 181 VRNSFLOESNGEERNITSPFSPGMYFEMIYCDVREKAVNGVHSLVKHFKELSS 240

QY 117 LRELIRSGSVOLYCVHS 133  
Db 241 IDTLEINGDIHLLEVR 257

RESULT 22  
ID ADO21123  
AC ADO21123;  
XX 18-NOV-2004 (first entry)  
XX Human diagnostic and therapeutic pprotein SEQ ID NO:3596.  
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX Homo sapiens.  
XX WO2004023973-A2.  
XX 25-MAR-2004.  
XX 12-SEP-2003; 2003WO-US028227.  
XX 12-SEP-2002; 2002US-0410259P.  
XX 12-SEP-2002; 2002US-0410260P.  
XX (INCY-) INCYTE CORP.  
XX

PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;  
XX  
XX WPI: 2004-329368/30.  
DR N-PSDB; ACM41999.  
DR  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
XX Claim 27; Page; 190pp; English.  
XX  
XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorders, developmental disorders, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 258 AA;  
Query Match 16.1%; Score 109.5; DB 8; Length 258;  
Best Local Similarity 25.5%; Pred. No. 8.4e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPQGVIIVRGLVQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 121 TQQLRLPFAARLNTPMGPGRTVVVKGEVANAKSFNVDLLAGSKDIALHLNPLNIKAF 180  
QY 57 DRTLAWISRWGQKLIAPFLYPPORFFVLLFQEGGLKALNGQGLGATSNMQOALEQ 116  
Db 181 VRNSFLQESWGGEERNITSPFSPGMYFEMIIYCDVREPKVAVNGVHSLEYKHFRELSS 240  
QY 117 LRELIRSGSVQLYCVHS 133  
Db 241 IDTLEINGDIHLLEVR 257  
RESULT 23  
ABM83348  
ID ABM83348 standard; protein; 258 AA.  
XX  
XX ABM83348;  
XX  
XX 18-NOV-2004 (first entry)  
DT  
DE Human diagnostic and therapeutic pprotein SEQ ID NO:3597.  
XX  
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX Homo sapiens.  
XX WO2004023973-A2.  
PN  
XX  
XX 25-MAR-2004.  
XX

PF 12-SEP-2003; 2003WO-US028227.  
XX  
PR 12-SEP-2002; 2002US-0410259P.  
PR 12-SEP-2002; 2002US-0410260P.  
XX (INCY-) INCYTE CORP.  
PA  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;  
XX  
XX WPI: 2004-329368/30.  
DR N-PSDB; ACM42000.  
DR  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
XX Claim 27; Page; 190pp; English.  
XX  
XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorders, developmental disorders, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 258 AA;  
Query Match 16.1%; Score 109.5; DB 8; Length 258;  
Best Local Similarity 25.5%; Pred. No. 8.4e-05;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPQGVIIVRGLVQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 121 TQQLRLPFAARLNTPMGPGRTVVVKGEVANAKSFNVDLLAGSKDIALHLNPLNIKAF 180  
QY 57 DRTLAWISRWGQKLIAPFLYPPORFFVLLFQEGGLKALNGQGLGATSNMQOALEQ 116  
Db 181 VRNSFLQESWGGEERNITSPFSPGMYFEMIIYCDVREPKVAVNGVHSLEYKHFRELSS 240  
QY 117 LRELIRSGSVQLYCVHS 133  
Db 241 IDTLEINGDIHLLEVR 257  
RESULT 24  
AA87403  
ID AA87403 standard; protein; 316 AA.  
XX  
XX AA87403;  
XX  
XX 03-JUL-2000 (first entry)  
DT  
DE Human PCTA-1 splice variant #1, SEQ ID NO:5.  
XX  
XX Prostate carcinoma antigen-1; PCTA-1; human; galectin-8 homologue;  
KW

KW splice variant; prostate cancer; biallelic marker; polymorphism;  
KW detection; diagnosis; prognosis; drug screening.  
XX  
XX Homo sapiens.  
XX WO9964590-A1.  
XX  
XX 16-DEC-1999.  
XX  
XX 04-JUN-1999; 99WO-IB001072.  
XX  
XX 05-JUN-1998; 98US-0088187P.  
XX  
XX 28-SEP-1998; 98US-0102324P.  
XX  
XX (GEST ) GENSET.  
XX  
XX Blumenfeld M, Bougueleret L, Chumakov I;  
XX  
XX WPI; 2000-338635/29.  
XX  
XX N-PSDB; AAA10226.  
XX  
XX Novel polynucleotide used to develop products for detecting  
XX susceptibility to, diagnosis, prognosis and therapy of prostate cancer.  
XX  
XX Claim 31; Page 329-330; 339pp; English.  
XX  
XX This sequence represents a splice variant (#1) of human prostate  
XX carcinoma tumour antigen-1 (PCTA-1). At least three PCTA-1 cDNAs  
XX (AAA10226, AAA10227, AAA10228) have been identified; the cDNA encoding  
XX this sequence comprises exons 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 out of the  
XX thirteen exons of the genomic PCTA-1 sequence (AAA10225). PCTA-1 is a  
XX human homologue of rat galectin-8, which can mediate both cell-cell and  
XX cell-matrix interactions. 125 biallelic markers were identified in the  
XX genomic sequence, with 40 biallelic markers being located in exonic  
XX regions. Six of these alter the amino acid sequence of a PCTA-1 protein.  
XX 43 markers were found in the 5' regulatory region, including 16 in the  
XX promoter, 39 in intronic sequences, and 3 in the 3' regulatory region.  
XX Alleles containing certain biallelic markers are associated with prostate  
XX cancer. These polymorphisms can be used as the basis of methods for  
XX determining genetic predisposition to prostate cancer, for diagnosis of  
XX this cancer and for the prognosis/detection of an eventual treatment  
XX response to therapeutic agents against prostate cancer. They can be used  
XX to design of drug screening protocols to provide accurate and efficient  
XX evaluation of therapeutic and side-effect potential of new or already  
XX existing medicaments, as well as for screening molecules which modulate  
XX or inhibit the expression of the PCTA-1 gene, preferably those active  
XX against prostate cancer  
XX  
XX Sequence 316 AA;

Query Match 16.1%; Score 109.5; DB 3; Length 316;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDQAHAAPVTLRASFA 56  
DB 179 TPQLSLPFAARLNTPMGRTVVVKGVEVNAKSFNVDDLAKGSKDIALHNLNRLNIKAF 238  
QY 57 DRTLAWISRWGQKKLISAPFLFYQRFVLLLFQEGGLKALNGQGLGATSMNQALEQ 116  
DB 239 VRNSFLOESWGEEERNITSPFPSPGMVFEMIIYCDVREFKAVNGVSHLSYKHKRFKELSS 298  
QY 117 LRELIRSGSVOLYCVHS 133  
DB 299 IDTLEINGDIHLLVEVS 315

RESULT 25  
AAB85030  
ID AAB85030 standard; protein; 316 AA.  
XX  
XX AAB85030;  
AC  
XX

06-AUG-2001 (first entry)  
XX  
XX Protein encoded by a PCTA-1 cDNA consisting of exons 0 to 9.  
XX  
XX BAP28; prostate; tumour; cancer; diagnostic; genetic analysis; PCTA-1.  
XX  
XX Homo sapiens.  
XX  
XX WO200100569-A2.  
XX  
XX 04-JAN-2001.  
XX  
XX 23-JUN-2000; 2000WO-IB001183.  
XX  
XX 25-JUN-1999; 99US-0141323P.  
XX  
XX 18-JAN-2000; 2000US-0176880P.  
XX  
XX (GEST ) GENSET.  
XX  
XX Barry C, Bougueleret L, Chumakov I, Cohen-Akenine A;  
XX  
XX WPI; 2001-367032/38.  
XX  
XX N-PSDB; AAP83912.  
XX  
XX New BAP28 polynucleotides and polypeptides overexpressed in prostate  
XX cancer cells for diagnosing prostate tumors, e.g. by hybridization or  
XX polymerase chain reaction assays.  
XX  
XX Example; Page 304-307; 349pp; English.  
XX  
XX The invention is directed to BAP28 polypeptides, BAP28 polynucleotide  
XX sequences and regulatory region located at the 3' and 5' ends of the  
XX BAP28 coding region. The BAP28 polypeptides can be expressed by standard  
XX recombinant methodology. BAP28 polynucleotides and polypeptides have been  
XX found to be over expressed in prostate tumour cells, therefore levels of  
XX BAP28 expression and/or activity may be assayed (e.g. by polymerase chain  
XX reaction (PCR)) to diagnose patient suffering from or susceptible to  
XX prostate cancer. Antibodies specific for the BAP28 polypeptides are  
XX useful as diagnostic reagents. Biallelic markers of the BAP28 gene are  
XX useful in genetic analysis. The present sequence represents a protein  
XX encoded by a first cDNA sequence of the PCTA-1 gene consisting of the  
XX exons 0 to 9. The coding strand of PCTA-1 gene is on the opposite of the  
XX coding strand of BAP28

Query Match 16.1%; Score 109.5; DB 4; Length 316;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
QY 2 SPRLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDQAHAAPVTLRASFA 56  
DB 179 TPQLSLPFAARLNTPMGRTVVVKGVEVNAKSFNVDDLAKGSKDIALHNLNRLNIKAF 238  
QY 57 DRTLAWISRWGQKKLISAPFLFYQRFVLLLFQEGGLKALNGQGLGATSMNQALEQ 116  
DB 239 VRNSFLOESWGEEERNITSPFPSPGMVFEMIIYCDVREFKAVNGVSHLSYKHKRFKELSS 298  
QY 117 LRELIRSGSVOLYCVHS 133  
DB 299 IDTLEINGDIHLLVEVS 315

RESULT 26  
AD057924  
ID AD057924 standard; protein; 316 AA.  
XX  
XX AD057924;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human galectin-8 polypeptide.  
DE  
XX



CC	or mRNA	
XX	Sequence 317 AA;	
SQ	Sequence 317 AA;	
	Query Match	16.1%; Score 109.5; DB 2; Length 317;
	Best Local Similarity	25.5%; Pred. No. 0.00011;
	Matches	35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;
QY	2 SPRLVPCSHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDOAAHAPVTLRASFA 56	
DB	180 TPQLSLFPAARLNTPMGPGRTVVVKGEVNNAKSFNVDDLAKGSKDIALHLNPLNLKAF 239	
QY	57 DRTLAWISRWGQKLLISAPFLFYPQRFVEVLLFQEGGLKALNGOGLGATSMNQALEQ 116	
DB	240 VRNSFLOESWGEERNITAPFPSPGMVFEMIIYCDVREFKAVNGVHSLYKHKRFKELSS 299	
QY	117 LRELIRSGSVOLYCVHS 133	
DB	300 IDTLEINGDIHLLEVR 316	
RESULT 28		
ADS34896		
ID	ADS34896 standard; protein; 317 AA.	
XX	ADS34896;	
AC		
DT	16-DEC-2004 (first entry)	
XX	Human autoimmune disease-related protein - SEQ ID 110.	
DE		
XX	single nucleotide polymorphism detection; SNP detection;	
KW	rheumatoid arthritis; type 1 diabetes; multiple sclerosis;	
KW	systemic lupus erythematosus; inflammatory bowel disease; psoriasis;	
KW	thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;	
KW	glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;	
KW	primary systemic vasculitis.	
OS	Homo sapiens.	
XX		
PN	WO2004083403-A2.	
XX		
PD	30-SEP-2004.	
XX		
PF	18-MAR-2004; 2004WO-US008461.	
XX		
PR	18-MAR-2003; 2003US-0455444P.	
PR	25-APR-2003; 2003US-0465241P.	
XX	(APPL-) APPLERA CORP.	
PA		
PI	Cargill M, Begovich AB, Alexander HC;	
XX		
DR	WPI; 2004-728480/71.	
DR	N-PSDB; ADS34813.	
XX		
PT	New isolated nucleic acid molecule comprises at least 8 contiguous	
PT	nucleotides where one of the nucleotides is a single nucleotide	
PT	polymorphism (SNP), useful for diagnosing or treating autoimmune	
PT	diseases, e.g. rheumatoid arthritis.	
XX		
PS	Claim 12; SEQ ID NO 110; 123pp; English.	
XX		
CC	The invention comprises amino acid and coding sequences containing	
CC	genetic polymorphisms associated with an altered risk of developing an	
CC	autoimmune disease (e.g. rheumatoid arthritis). The invention further	
CC	comprises a method of identifying an individual that has an altered risk	
CC	of developing an autoimmune disease, comprising detecting a single	
CC	nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA	
CC	and protein sequences of the invention are useful for diagnosing and	
CC	treating autoimmune diseases, such as: rheumatoid arthritis, type 1	
CC	diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory	
CC	bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious	

CC	anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,	
CC	myocarditis, Sjogren's disease, or primary systemic vasculitis. The	
CC	present amino acid sequence represents a human autoimmune disease-related	
CC	protein of the invention. NOTE: The present sequence is not shown in the	
CC	specification, but has been retrieved from the WIPO website.	
XX		
SQ	Sequence 317 AA;	
	Query Match	16.1%; Score 109.5; DB 8; Length 317;
	Best Local Similarity	25.5%; Pred. No. 0.00011;
	Matches	35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;
QY	2 SPRLVPCSHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDOAAHAPVTLRASFA 56	
DB	180 TPQLRUPFAARLNTPMGPGRTVVVKGEVNNAKSFNVDDLAKGSKDIALHLNPLNLKAF 239	
QY	57 DRTLAWISRWGQKLLISAPFLFYPQRFVEVLLFQEGGLKALNGOGLGATSMNQALEQ 116	
DB	240 VRNSFLOESWGEERNITSPFPSPGMVFEMIIYCDVREFKAVNGVHSLYKHKRFKELSS 299	
QY	117 LRELIRSGSVOLYCVHS 133	
DB	300 IDTLEINGDIHLLEVR 316	
RESULT 29		
ADS34895		
ID	ADS34895 standard; protein; 317 AA.	
XX	ADS34895;	
AC		
DT	16-DEC-2004 (first entry)	
XX	Human autoimmune disease-related protein - SEQ ID 109.	
DE		
XX	single nucleotide polymorphism detection; SNP detection;	
KW	rheumatoid arthritis; type 1 diabetes; multiple sclerosis;	
KW	systemic lupus erythematosus; inflammatory bowel disease; psoriasis;	
KW	thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;	
KW	glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;	
KW	primary systemic vasculitis.	
OS	Homo sapiens.	
XX		
PN	WO2004083403-A2.	
XX		
PD	30-SEP-2004.	
XX		
PF	18-MAR-2004; 2004WO-US008461.	
XX		
PR	18-MAR-2003; 2003US-0455444P.	
PR	25-APR-2003; 2003US-0465241P.	
XX	(APPL-) APPLERA CORP.	
PA		
PI	Cargill M, Begovich AB, Alexander HC;	
XX		
DR	WPI; 2004-728480/71.	
DR	N-PSDB; ADS34812.	
XX		
PT	New isolated nucleic acid molecule comprises at least 8 contiguous	
PT	nucleotides where one of the nucleotides is a single nucleotide	
PT	polymorphism (SNP), useful for diagnosing or treating autoimmune	
PT	diseases, e.g. rheumatoid arthritis.	
XX		
PS	Claim 12; SEQ ID NO 109; 123pp; English.	
XX		
CC	The invention comprises amino acid and coding sequences containing	
CC	genetic polymorphisms associated with an altered risk of developing an	
CC	autoimmune disease (e.g. rheumatoid arthritis). The invention further	
CC	comprises a method of identifying an individual that has an altered risk	
CC	of developing an autoimmune disease, comprising detecting a single	
CC	nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA	
CC	and protein sequences of the invention are useful for diagnosing and	
CC	treating autoimmune diseases, such as: rheumatoid arthritis, type 1	
CC	diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory	
CC	bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious	

CC and protein sequences of the invention are useful for diagnosing and  
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1  
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory  
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious  
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,  
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The  
CC present amino acid sequence represents a human autoimmune disease-related  
CC protein of the invention. NOTE: The present sequence is not shown in the  
CC specification, but has been retrieved from the WIPO website.  
XX  
SQ Sequence 317 AA;  
  
Query Match 16.1%; Score 109.5; DB 8; Length 317;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
  
QY 2 SPRLEVPCHALPQGLSPQGVIIIRGLVLQEPKHPTVSL-----RDQAAHAPVTLRASFA 56  
Db 180 TPQLRLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLGKSKDIALHLNPLNIKAF 239  
  
QY 57 DRTLAWISRWGQKKLISAPFLYPQRPFFVLLLFQGGKLLALNGOGLGATSMNQQALEQ 116  
Db 240 VRNSFLQESWGEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLKYHRFKELSS 299  
  
QY 117 LRELISGVSQLYCVHS 133  
Db 300 IDTLEINGDIHLEVRSS 316  
  
RESULT 30  
ABP65200  
ID ABP65200 standard; protein; 318 AA.  
XX  
AC ABP65200;  
XX  
DT 12-NOV-2002 (first entry)  
XX  
DE Hypoxia-regulated protein #74.  
XX  
KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;  
KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;  
KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;  
KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;  
KW preeclampsia; atherosclerosis; inflammatory condition; wound healing;  
KW inflammation; erythropoiesis; hair loss; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200246465-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 10-DEC-2001; 2001WO-GB005458.  
XX  
PR 08-DEC-2000; 2000GB-00030076.  
PR 08-FEB-2001; 2001GB-00003156.  
PR 25-OCT-2001; 2001GB-00025666.  
XX  
PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX  
PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;  
PI Rayner WN;  
XX  
DR WPI; 2002-627238/67.  
XX  
PT Identifying a gene involved in disease for treating hypoxia-regulated  
PT conditions, comprises comparing the transcriptome/proteome of two cell  
PT types under different conditions and identifying a differentially  
PT regulated gene.  
XX  
PS Claim 35; Page 389; 538pp; English.  
XX  
CC The present invention relates to methods for identifying genes and

CC proteins that are implicated in a specific disease or physiological  
CC condition. The method comprises comparing the transcriptome/proteome of a  
CC specialised cell type implicated in a disease or condition with that of a  
CC second specialised cell type, under two experimental conditions, and  
CC identifying a gene that is differentially regulated in the two  
CC specialised cell types under experimental conditions. ABV7873-ABV78116  
CC and ABP65061-ABP65257 were identified using the methods of the invention.  
CC The coding sequences and proteins are useful for treating a disease in a  
CC patient, for manufacture of a medicament for treating hypoxia-regulated  
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,  
CC biological response to hypoxia conditions, or hypoxic-associated  
CC pathology in a patient. The coding sequences and proteins are also useful  
CC for monitoring the therapeutic treatment of a disease or physiological  
CC condition, such as cancer, ischaemic conditions, reperfusion injury,  
CC retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory  
CC conditions, wound healing, inflammation, erythropoiesis or hair loss  
XX  
SQ Sequence 318 AA;  
  
Query Match 16.1%; Score 109.5; DB 5; Length 318;  
Best Local Similarity 25.5%; Pred. No. 0.00011;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
  
QY 2 SPRLEVPCHALPQGLSPQGVIIIRGLVLQEPKHPTVSL-----RDQAAHAPVTLRASFA 56  
Db 181 TPQLRLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLGKSKDIALHLNPLNIKAF 240  
  
QY 57 DRTLAWISRWGQKKLISAPFLYPQRPFFVLLLFQGGKLLALNGOGLGATSMNQQALEQ 116  
Db 241 VRNSFLQESWGEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLKYHRFKELSS 300  
  
QY 117 LRELISGVSQLYCVHS 133  
Db 301 IDTLEINGDIHLEVRSS 317  
  
Search completed: October 17, 2005, 08:39:25  
Job time : 170 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:35:58 ; Search time 42 Seconds  
(without alignments)

236.389 Million cell updates/sec

Title: HJACE54

Perfect score: 682

Sequence: 1 msrlevpcshalpgqlspg.....leqlrelriagsvqlycvhs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	671	98.4	133	4	US-09-557-170A-2
2	671	98.4	275	4	US-09-557-170A-25
3	671	98.4	296	4	US-09-557-170A-27
4	671	98.4	336	3	US-09-131-648-1
5	119	17.4	145	2	US-08-788-584-5
6	119	17.4	145	3	US-08-946-914-12
7	119	17.4	145	4	US-09-656-450-12
8	119	17.4	145	4	US-09-557-170A-3
9	109.5	16.1	316	4	US-09-326-402C-5
10	109.5	16.1	317	3	US-09-326-402C-15
11	109.5	16.1	317	4	US-08-946-914-6
12	109.5	16.1	317	4	US-09-656-450-6
13	109.5	16.1	318	4	US-09-557-170A-4
14	108.5	15.9	316	3	US-09-131-648-5
15	108.5	15.9	317	4	US-08-875-553D-30
16	106.5	15.6	368	4	US-09-326-402C-7
17	98	14.4	322	4	US-09-559-023-2
18	96.5	14.1	358	4	US-09-326-402C-6
19	96.5	14.1	358	4	US-09-326-402C-16
20	95.5	14.0	135	4	US-09-326-402C-20
21	95.5	14.0	136	3	US-08-946-914-13
22	95.5	14.0	136	4	US-09-154-750A-79
23	95.5	14.0	136	4	US-09-656-450-13
24	94	13.8	315	4	US-09-326-402C-17
25	91.5	13.4	316	4	US-09-326-402C-9
26	90.5	13.3	139	4	US-09-513-999C-6729
27	88.5	13.0	264	1	US-08-562-311-4
28	87.5	12.8	316	2	US-08-728-521-3
29	87.5	12.8	316	2	US-08-647-960-2
30	87.5	12.8	316	3	US-08-946-914-15
31	87.5	12.8	316	3	US-08-946-914-17
32	87.5	12.8	316	3	US-09-212-146-3
33	87.5	12.8	316	4	US-09-656-450-15
34	87.5	12.8	316	4	US-09-656-450-17
35	86.5	12.7	264	2	US-08-728-521-1
36	86.5	12.7	264	3	US-09-212-146-1
37	84	12.3	324	3	US-08-946-914-11
38	84	12.3	324	4	US-09-656-450-11
39	82.5	12.1	250	1	US-08-562-311-2
40	81.5	12.0	323	1	US-08-469-667-16
41	81.5	12.0	323	3	US-08-946-914-2
42	81.5	12.0	323	3	US-09-224-110-16
43	81.5	12.0	323	4	US-09-656-450-2
44	81.5	12.0	323	4	US-09-326-402C-21
45	81.5	12.0	323	4	US-09-988-292A-16
46	81.5	12.0	323	5	PCT-US95-07289-16
47	81.5	12.0	341	4	US-09-949-016-7251
48	81	11.9	273	4	US-09-902-540-9846
49	81	11.9	737	3	US-09-511-625B-2
50	81	11.9	819	3	US-09-511-625B-4
51	81	11.9	847	1	US-08-276-099A-2
52	81	11.9	847	1	US-08-781-890-2
53	81	11.9	847	3	US-09-087-465-12
54	81	11.9	847	3	US-09-511-625B-68
55	81	11.9	847	4	US-09-972-800A-10
56	80	11.7	145	2	US-08-788-584-1
57	80	11.7	200	3	US-08-946-914-8
58	80	11.7	200	4	US-09-656-450-8
59	78.5	11.5	262	3	US-08-946-914-14
60	78.5	11.5	262	4	US-09-656-450-14
61	78.5	11.5	282	4	US-09-489-039A-12241
62	77.5	11.4	143	4	US-09-877-790-1
63	77.5	11.4	249	4	US-09-538-092-951
64	77.5	11.4	250	3	US-08-946-914-10
65	77.5	11.4	250	4	US-09-656-450-10
66	77.5	11.4	250	4	US-09-919-039-298
67	77.5	11.4	250	4	US-09-877-790-2
68	77.5	11.4	251	4	US-09-949-016-7560
69	76	11.1	311	3	US-08-946-914-4
70	76	11.1	311	4	US-09-656-450-4
71	76	11.1	355	4	US-09-326-402C-18
72	76	11.1	378	4	US-09-854-133-439
73	75	11.0	149	2	US-08-788-584-3
74	75	11.0	343	4	US-09-252-991A-27631
75	75	11.0	368	4	US-09-902-540-13302
76	75	11.0	369	4	US-09-543-681A-7514
77	74.5	10.9	146	2	US-08-647-960-8
78	73	10.7	323	4	US-09-326-402C-19
79	71.5	10.5	332	4	US-09-135-121B-7
80	71	10.4	134	4	US-09-326-402C-14
81	71	10.4	135	3	US-08-946-914-16
82	71	10.4	135	3	US-08-050-259B-20
83	71	10.4	135	3	US-09-489-292-2
84	71	10.4	135	4	US-08-719-579-2
85	71	10.4	135	4	US-09-656-450-16
86	71	10.4	135	4	US-09-919-172-91
87	71	10.4	135	4	US-09-976-594-628
88	71	10.4	135	4	US-09-919-039-242
89	71	10.4	135	4	US-09-919-497-81
90	71	10.4	149	4	US-09-949-016-10945
91	71	10.4	422	4	US-09-270-767-42572
92	71	10.4	617	4	US-09-252-991A-21113
93	70	10.3	303	4	US-08-583-110-4995
94	70	10.3	347	4	US-09-107-433-5171
95	70	10.3	373	4	US-09-134-000C-5355
96	70	10.3	545	4	US-09-949-016-10223
97	69.5	10.2	206	4	US-09-180-109A-15
98	69.5	10.2	206	4	US-09-180-109A-18
99	69	10.1	352	3	US-09-180-109A-18
100	68.5	10.0	320	4	US-09-252-991A-23385

Sequence 3, Appli  
Sequence 2, Appli  
Sequence 15, Appli  
Sequence 17, Appli  
Sequence 3, Appli  
Sequence 15, Appli  
Sequence 17, Appli  
Sequence 1, Appli  
Sequence 11, Appli  
Sequence 11, Appli  
Sequence 16, Appli  
Sequence 16, Appli  
Sequence 2, Appli  
Sequence 16, Appli  
Sequence 21, Appli  
Sequence 2, Appli  
Sequence 16, Appli  
Sequence 15, Appli  
Sequence 7851, Ap  
Sequence 9846, Ap  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 12, Appli  
Sequence 68, Appli  
Sequence 10, Appli  
Sequence 1, Appli  
Sequence 8, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 12241, A  
Sequence 1, Appli  
Sequence 951, App  
Sequence 10, Appli  
Sequence 10, Appli  
Sequence 298, App  
Sequence 2, Appli  
Sequence 7560, Ap  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 18, Appli  
Sequence 439, App  
Sequence 3, Appli  
Sequence 27631, A  
Sequence 13302, A  
Sequence 7514, Ap  
Sequence 8, Appli  
Sequence 19, Appli  
Sequence 7, Appli  
Sequence 14, Appli  
Sequence 16, Appli  
Sequence 20, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 16, Appli  
Sequence 628, App  
Sequence 242, App  
Sequence 81, Appli  
Sequence 10945, A  
Sequence 42572, A  
Sequence 21113, A  
Sequence 4995, Ap  
Sequence 5171, Ap  
Sequence 5355, Ap  
Sequence 10223, A  
Sequence 15, Appli  
Sequence 18, Appli  
Sequence 4121, Ap  
Sequence 23385, A

## ALIGNMENTS

```
RESULT 1
US-09-557-170A-2
; Sequence 2, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-2

Query Match          98.4%; Score 671; DB 4; Length 133;
Best Local Similarity 99.2%; Pred. No. 1.3e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
QY 61 AWISRWGQKKLIASAPLFYPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 61 AWISRWGQKKLIASAPLFYPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
QY 121 RISGSVQLYCVHS 133
Db 121 RISGSVQLYCVHS 133

RESULT 2
US-09-557-170A-25
; Sequence 25, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
```

```
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-25
```

```
Query Match          98.4%; Score 671; DB 4; Length 275;
Best Local Similarity 99.2%; Pred. No. 3.5e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 143 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 202
QY 61 AWISRWGQKKLIASAPLFYPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 203 AWISRWGQKKLIASAPLFYPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 262
QY 121 RISGSVQLYCVHS 133
Db 263 RISGSVQLYCVHS 275
```

## RESULT 3

```
US-09-557-170A-27
; Sequence 27, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-27
```

```
Query Match          98.4%; Score 671; DB 4; Length 296;
Best Local Similarity 99.2%; Pred. No. 3.9e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 60
Db 164 MSPRLVPCSHALPQGLSPGQVIIVRGLVLOEPKHPFTVSLRDOAAHAPVTLRASPADRTL 223
QY 61 AWISRWGQKKLIASAPLFYPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 120
Db 224 AWISRWGQKKLIASAPLFYPQRFVFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLREL 283
QY 121 RISGSVQLYCVHS 133
Db 284 RISGSVQLYCVHS 296
```



```

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0192 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 727176
US-08-788-584-5

Query Match 17.4%; Score 119; DB 2; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

Qy 3 PRLEVPCHALPQGLSPGQVIVRGLVQPKHPTVSLR---DQAAHAPVTLRASFA DRT 59
Db 11 PNLAVPFTSIENGLVPSKSIIVISGVLSDAKRFQINLRGCGDIAFHLPNPFDENAVVRN 70
Qy 60 LAWSRWG-QKKLISAPFLFYQRFPEVLLVFOEGGLKALNGQGLGATSMNQALEOLR 118
Db 71 TQINNSWGFERSLPQSMFPRSGQRFVWILCEGHCFKVAVDQGHICYSYHRLMNLDPIN 130
Qy 119 ELRISGVOL 128
Db 131 TLEVAGDIQL 140

RESULT 6
US-08-946-914-12
Sequence 12, Application US/08946914
Patent No. 6027916
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Gentz, Reiner L.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
STREET: 1100 New York Ave., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,914
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,093
FILING DATE: 09-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0560001/EKS/SGW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0192 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 727176
US-08-788-584-5

Query Match 98.4%; Score 671; DB 3; Length 336;
Best Local Similarity 99.2%; Pred. No. 4.6e-76;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSRLPVPCHALPQGLSPGQVIVRGLVQPKHPTVSLRDOAAHAPVTLRASFA DRTL 60
Db 204 MSRLPVPCHALPQGLSPGQVIVRGLVQPKHPTVSLRDOAAHAPVTLRASFA DRTL 263
Qy 61 AWISRWGQKKLISAPFLFYQRFPEVLLVFOEGGLKALNGQGLGATSMNQALEOLREL 120
Db 264 AWISRWGQKKLISAPFLFYQRFPEVLLVFOEGGLKALNGQGLGATSMNQALEOLREL 323
Qy 121 RISGSVOLYCVHS 133
Db 324 RISGSVOLYCVHS 336

RESULT 5
US-08-788-584-5
Sequence 5, Application US/08788584
Patent No. 5837493
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
APPLICANT: Bandman, Olga
APPLICANT: Hawkins, Phillip R.
APPLICANT: Pettithory, Joanne R.
TITLE OF INVENTION: NOVEL HUMAN GALECTINS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,584
FILING DATE: Filed Herewith
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

```

```
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-946-914-12

Query Match      17.4%; Score 119; DB 3; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPQGLSPGQVVIIVRGLVLOEPKHPVTSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFFTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYQRPFEVLLLFQEGGLKALNGQGLGATSMNQQALEQLR 118
Db 71 TQINNWSGPEERSLPGSMPPFSRGQRFVSWILCEGHCFKVAVDQGHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 7
US-09-656-450-12
; Sequence 12, Application US/09656450
; Patent No. 6468768
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560003
; CURRENT APPLICATION NUMBER: US/09/656,450
; CURRENT FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: US 60/028,093
; PRIOR FILING DATE: 1996-10-09
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Rat
US-09-656-450-12

Query Match      17.4%; Score 119; DB 4; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPQGLSPGQVVIIVRGLVLOEPKHPVTSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFFTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYQRPFEVLLLFQEGGLKALNGQGLGATSMNQQALEQLR 118
Db 71 TQINNWSGPEERSLPGSMPPFSRGQRFVSWILCEGHCFKVAVDQGHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 8
US-09-557-170A-3
; Sequence 3, Application US/09557170A
; Patent No. 6605699
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
```

```
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/09/557,170A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-170A-3

Query Match      17.4%; Score 119; DB 4; Length 145;
Best Local Similarity 26.9%; Pred. No. 5.6e-07;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;

QY 3 PRLEVPCHALPQGLSPGQVVIIVRGLVLOEPKHPVTSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFFTSIPNGLYPSKSIIVISGVVLSDAKRFQINLRCGGDIAFHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYQRPFEVLLLFQEGGLKALNGQGLGATSMNQQALEQLR 118
Db 71 TQINNWSGPEERSLPGSMPPFSRGQRFVSWILCEGHCFKVAVDQGHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 9
US-09-326-402C-5
; Sequence 5, Application US/09326402C
; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-T112XC1
; CURRENT APPLICATION NUMBER: US/09/326,402C
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 18
; OTHER INFORMATION: 5-2-162 : polymorphic amino acid Tyr or Phe
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 35
; OTHER INFORMATION: 5-2-213 : polymorphic amino acid Cys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 55
```

```
; OTHER INFORMATION: 5-3-84 : polymorphic amino acid Val or Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 183
; OTHER INFORMATION: 5-7-195 : polymorphic amino acid Ser or Arg
US-09-326-402C-5

Query Match          16.1%; Score 109.5; DB 4; Length 316;
Best Local Similarity 25.5%; Pred. No. 2.6e-05;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLVPCSHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDOAAHAPVTLRASFA 56
Db 179 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGSKDIALHLPRLNIKAF 238
QY 57 DRTLAWISRWGQKKLISAPFLYPQRFVFLVLLFQEGGLKALNGQGLGATSMNQALEQ 116
Db 239 VRNSFLQESWGEERNTSPFSPGMVFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSS 298
QY 117 LRELIRISGSVOLYCVHS 133
Db 299 IDTLEINGDIHLLVRS 315

RESULT 10
US-09-326-402C-15
; Sequence 15, Application US/09326402C
; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-T12XC1
; CURRENT APPLICATION NUMBER: US/09/326,402C
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(316)
; OTHER INFORMATION: amino acid sequence of PCTA
US-09-326-402C-15

Query Match          16.1%; Score 109.5; DB 4; Length 316;
Best Local Similarity 25.5%; Pred. No. 2.6e-05;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLVPCSHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDOAAHAPVTLRASFA 56
Db 179 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGSKDIALHLPRLNIKAF 238
QY 57 DRTLAWISRWGQKKLISAPFLYPQRFVFLVLLFQEGGLKALNGQGLGATSMNQALEQ 116
Db 239 VRNSFLQESWGEERNTSPFSPGMVFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSS 298
QY 117 LRELIRISGSVOLYCVHS 133
Db 299 IDTLEINGDIHLLVRS 315

RESULT 11
US-08-946-914-6
; Sequence 6, Application US/08946914
; Patent No. 6027916
```

```
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,914
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,093
; FILING DATE: 09-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION/DOCKET NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0560001/EKS/SGW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-946-914-6

Query Match          16.1%; Score 109.5; DB 3; Length 317;
Best Local Similarity 25.5%; Pred. No. 2.6e-05;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY 2 SPRLVPCSHALPQGLSPGVIIIRGLVLPQPKHPTVSL-----RDOAAHAPVTLRASFA 56
Db 180 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGSKDIALHLPRLNIKAF 239
QY 57 DRTLAWISRWGQKKLISAPFLYPQRFVFLVLLFQEGGLKALNGQGLGATSMNQALEQ 116
Db 240 VRNSFLQESWGEERNTAFPFSPGMVFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSS 299
QY 117 LRELIRISGSVOLYCVHS 133
Db 300 IDTLEINGDIHLLVRS 316

RESULT 12
US-09-656-450-6
; Sequence 6, Application US/09656450
; Patent No. 6468768
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560003
; CURRENT APPLICATION NUMBER: US/09/656,450
; CURRENT FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
```



QY 117 LRELRISSGVOLCYVHS 133  
:  
Db 300 IDTLEINGDIHLLEVRS 316

RESULT 16  
US-09-326-402C-7  
; Sequence 7, Application US/09326402C  
; Patent No. 6759192  
; GENERAL INFORMATION:  
; APPLICANT: Blumefeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)  
; FILE REFERENCE: GEN-T112XC1  
; CURRENT APPLICATION NUMBER: US/09/326.402C  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,187  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/102,324  
; PRIOR FILING DATE: 1998-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 18  
; OTHER INFORMATION: 5-2-162 : polymorphic amino acid Tyr or Phe  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 35  
; OTHER INFORMATION: 5-2-213 : polymorphic amino acid Cys or Arg  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 55  
; OTHER INFORMATION: 5-3-84 : polymorphic amino acid Val or Met  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 183  
; OTHER INFORMATION: 5-7-195 : polymorphic amino acid Ser or Arg  
US-09-326-402C-7

Query Match 15.6%; Score 106.5; DB 4; Length 368;  
Best Local Similarity 25.0%; Pred. No. 7.5e-05;  
Matches 33; Conservative 23; Mismatches 71; Indels 5; Gaps 1;

QY 2 SPRLEVPCSHALPQGLSPGVIIIVRGLVLQPDPKHTVSL-----RDQAHAHPVTTRASF 56  
:  
Db 179 TPQLSFPFAARLNTFPMGPGRTVVVKGEVNANAKSFNDLLAGKSKDIALHLPRLNIKAF 238  
:  
QY 57 DRTLAWISRWOKKLISAPFLFYQRPFEVLLLFOEGLKUALNGOGLGATSMNQALEQ 116  
:  
Db 239 VRNSFLQBSWGEERNITSPFPSPGYEMIIYCDVREFKVAVNGVHSLEYKHFRFKELSS 298  
:  
QY 117 LRELRISSGVOL 128  
:  
Db 299 IDTLEINGDIHL 310

RESULT 17  
US-09-559-023-2  
; Sequence 2, Application US/09559023  
; Patent No. 6551796  
; GENERAL INFORMATION:  
; APPLICANT: Abramson, Ruth  
; APPLICANT: Leal-Pinto, Edgar  
; APPLICANT: Lipkowitz, Michael  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING URATE TRANSPORTER  
; FILE REFERENCE: 070165.0574

```

; LOCATION: 211
; OTHER INFORMATION: 5-202-117 : polymorphic amino acid Leu or Stop
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 225
; OTHER INFORMATION: 5-7-195 : polymorphic amino acid Ser or Arg
US-09-326-402C-6

Query Match          14.1%; Score 96.5; DB 4; Length 358;
Best Local Similarity 25.0%; Pred. No. 0.0013;
Matches 33; Conservative 21; Mismatches 73; Indels 5; Gaps 1;

QY 7 VPCSHALPQGLSPGQVIVIRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFAADRTL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 LPFAARLNTPMGRTVVVKGEVNAKSNFVLLAGKSKDIALHLNPRINIKAFVNRNF 285

QY 62 WISRWGQKLIAPFLFYQRFPEVILLFQEGGLKALNGOGLGATSMNQQALEQLREL 121
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 286 LQESWGEERNITSPFPSPGMYFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSSIDTLE 345

QY 122 IGSVOLYCVHS 133
Db 346 INGDHLLLEVR 357

RESULT 19
US-09-326-402C-16
; Sequence 16, Application US/09326402C
; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-T112XC1
; CURRENT APPLICATION NUMBER: US/09/326,402C
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(358)
; OTHER INFORMATION: amino acid sequence of PCTA.var
US-09-326-402C-16

Query Match          14.1%; Score 96.5; DB 4; Length 358;
Best Local Similarity 25.0%; Pred. No. 0.0013;
Matches 33; Conservative 21; Mismatches 73; Indels 5; Gaps 1;

QY 7 VPCSHALPQGLSPGQVIVIRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFAADRTL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 LPFAARLNTPMGRTVVVKGEVNAKSNFVLLAGKSKDIALHLNPRINIKAFVNRNF 285

QY 62 WISRWGQKLIAPFLFYQRFPEVILLFQEGGLKALNGOGLGATSMNQQALEQLREL 121
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 286 LQESWGEERNITSPFPSPGMYFEMIIYCDVREFKVAVNGVHSLEYKHKRFKELSSIDTLE 345

QY 122 IGSVOLYCVHS 133
Db 346 INGDHLLLEVR 357

RESULT 20
US-09-326-402C-20
; Sequence 20, Application US/09326402C
```

```

; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-T112XC1
; CURRENT APPLICATION NUMBER: US/09/326,402C
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(135)
; OTHER INFORMATION: amino acid sequence of leg7
US-09-326-402C-20

Query Match          14.0%; Score 95.5; DB 4; Length 135;
Best Local Similarity 28.0%; Pred. No. 0.00047;
Matches 37; Conservative 25; Mismatches 55; Indels 15; Gaps 6;

QY 7 VPCSHALPQGLSPGQVIVIRGLVLPQPKHPTVSL---RDQAAHAPVTLRASFAADRTL 63
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 3 VPHKSSLPEGIRPGTVLRIRGLVPPNASFVHVNLLCGEEQGDAAALHFNPRLDTSEWFN 62

QY 64 SR-----WG-QKKLISAPFLFYQRFPEVILLFQEGGLKALNGOGLGATSMNQQALEQ 116
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 63 SKEQSGWGREERGPGVPF---QRQPFVLLIADSDGFK-AVVGDAQYVHFHRLPLAR 117

QY 117 LRELIRSGSVQL 128
Db 118 VRLVEVGDDVQL 129

RESULT 21
US-08-946-914-13
; Sequence 13, Application US/08946914
; Patent No. 6027916
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,914
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,093
; FILING DATE: 09-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
```



QY 4 RLEVPCHALPQGLSPQGVIIVRGLVQLQEPKHTVSL-----RDQAAHAPVTLRASFADR 58  
Db 181 QLSLPEARLNASMGRTVWIKGEVNTNARSFVNDLVAGKTRDIALHLNPLNKAFAVR-R 239  
QY 59 TLAWISRWGQKKLISAPFLFYPPQRFVEVLLFOEGGLKIALNQGLGATSMNQALEQLR 118  
Db 240 NSFLQDANGEEERNITCFPPSSGMFYEMIIYCDVREFKVAINGVHSLEYKHKRPFKDLSSID 299  
QY 119 ELRISGSVQLYCVHS 133  
Db 300 TLSVDGDIRLLDVRS 314  
RESULT 25  
US-09-326-402C-9  
; Sequence 9, Application US/09326402C  
; Patent No. 6759192  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)  
; FILE REFERENCE: GEN-T112XC1  
; CURRENT APPLICATION NUMBER: US/09/326,402C  
; CURRENT FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,187  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/102,324  
; PRIOR FILING DATE: 1998-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 9  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-326-402C-9  
Query Match 13.4%; Score 91.5; DB 4; Length 316;  
Best Local Similarity 23.0%; Pred. No. 0.0046;  
Matches 31; Conservative 23; Mismatches 76; Indels 5; Gaps 1;  
QY 4 RLEVPCHALPQGLSPQGVIIVRGLVQLQEPKHTVSL-----RDQAAHAPVTLRASFADR 58  
Db 181 QLSLPEARLNASMGRTVWIKGEVNTNARSFVNDLVAGKTRDIALHLNPLNKAFAVR 240  
QY 59 TLAWISRWGQKKLISAPFLFYPPQRFVEVLLFOEGGLKIALNQGLGATSMNQALEQLR 118  
Db 241 NSFLQDANGEEERNITCFPPSSGMFYEMIIYCDVREFKVAINGVHSLEYKHKRPFKDLSSID 300  
QY 119 ELRISGSVQLYCVHS 133  
Db 301 TLSVDGDIRLLDVRS 315  
RESULT 26  
US-09-513-999C-6729  
; Sequence 6729, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.psm  
; SEQ ID NO 6729

; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6729  
Query Match 13.3%; Score 90.5; DB 4; Length 139;  
Best Local Similarity 33.7%; Pred. No. 0.0021;  
Matches 31; Conservative 15; Mismatches 35; Indels 11; Gaps 5;  
QY 3 RLEVP-CHALPQGLSPQGVIIVRGLVQLQEPKHTVSL-----RDQAAHAPVTLRASFA 56  
Db 34 PRLIVPFCGH-IKGWRPGKVKLVNMGIVDLNPESFAISLTGDSBPPADVAIELKAVFT 92  
QY 57 DRTL---AMIS-RWGOKKLISAPFLFYPPQRF 84  
Db 93 DRQLLNSCISGERGEEQSAIFYFFPIPDQPF 124  
RESULT 27  
US-08-562-311-4  
; Sequence 4, Application US/08562311  
; Patent No. 5801002  
; GENERAL INFORMATION:  
; APPLICANT: RAZ, AVRAHAM  
; TITLE OF INVENTION: A METHOD OF DETERMINING THE PROBABILITY  
; TITLE OF INVENTION: OF METASTASIS IN A CELL SAMPLE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dykema Gossett  
; STREET: STE 505 N. Woodward  
; CITY: Bloomfield Hills  
; STATE: MI  
; COUNTRY: U.S.  
; ZIP: 48304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/562,311  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION: 435  
; APPLICATION NUMBER: US/08/188,225  
; FILING DATE:  
; APPLICATION NUMBER: US 07/681,242  
; FILING DATE: 04-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/294,249  
; FILING DATE: 01-JUN-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KELLY, ROBERT L.  
; REGISTRATION NUMBER: 31,843  
; REFERENCE/DOCKET NUMBER: 61,686-  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 810-540-0849  
; TELEFAX: 810-540-0763  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-562-311-4  
Query Match 13.0%; Score 88.5; DB 1; Length 264;  
Best Local Similarity 26.1%; Pred. No. 0.0089;  
Matches 35; Conservative 21; Mismatches 63; Indels 15; Gaps 4;  
QY 5 LEVPCSHALPQGLSPQGVIIVRGLVQLQEPKHTVSLR---DQAAHAPVTLRASFA 61  
Db 128 LTVPDLPLPGGLMPRLITMGTVKPNANRIVLDFRGNDAFH----FNPRFNNRR 183





Search completed: October 17, 2005, 08:43:38  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:36:43 ; Search time 168 Seconds

(without alignments)  
329.991 Million cell updates/sec

Title: HJACE54

Perfect score: 682

Sequence: 1 msp1evpcshalpglspg.....leglreirsgsvqlcyvhs 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1860064 seqs, 416830855 residues

Total number of hits satisfying chosen parameters: 1860064

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*

20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	671	98.4	133 14	US-10-156-136-29
2	671	98.4	133 15	US-10-455-366-2
3	671	98.4	133 20	US-11-041-419-29
4	671	98.4	275 15	US-10-455-366-25
5	671	98.4	296 15	US-10-455-366-27
6	671	98.4	336 9	US-09-747-804-1
7	120.5	17.7	196 9	US-09-768-826-55
8	120.5	17.7	196 16	US-10-874-484-55
9	120.5	17.7	318 15	US-10-151-166-4
10	119	17.4	145 9	US-09-728-479-8
11	119	17.4	145 9	US-09-894-526-5
Sequence 29, Appli				
Sequence 2, Appli				
Sequence 29, Appli				
Sequence 25, Appli				
Sequence 27, Appli				
Sequence 1, Appli				
Sequence 55, Appli				
Sequence 55, Appli				
Sequence 4, Appli				
Sequence 8, Appli				
Sequence 5, Appli				

12	119	17.4	145	9	US-09-263-689-12	Sequence 12, Appli
13	119	17.4	145	14	US-10-235-674-12	Sequence 12, Appli
14	119	17.4	145	14	US-10-156-136-42	Sequence 42, Appli
15	119	17.4	145	15	US-10-455-366-3	Sequence 3, Appli
16	119	17.4	145	20	US-11-041-419-42	Sequence 42, Appli
17	109.5	16.1	158	17	US-10-491-213-9	Sequence 9, Appli
18	109.5	16.1	316	16	US-10-856-888-5	Sequence 5, Appli
19	109.5	16.1	316	16	US-10-856-888-15	Sequence 15, Appli
20	109.5	16.1	317	9	US-09-263-689-6	Sequence 6, Appli
21	109.5	16.1	317	14	US-10-235-674-6	Sequence 6, Appli
22	109.5	16.1	318	14	US-10-156-136-43	Sequence 43, Appli
23	109.5	16.1	318	14	US-10-376-133-19	Sequence 19, Appli
24	109.5	16.1	318	15	US-10-170-385-365	Sequence 365, App
25	109.5	16.1	318	15	US-10-455-366-4	Sequence 4, Appli
26	109.5	16.1	318	20	US-11-041-419-43	Sequence 43, Appli
27	108.5	15.9	316	9	US-09-747-804-5	Sequence 5, Appli
28	108.5	15.9	316	14	US-10-177-293-354	Sequence 354, App
29	108.5	15.9	316	17	US-10-482-029-134	Sequence 134, App
30	108.5	15.9	317	10	US-09-948-227-6	Sequence 6, Appli
31	108.5	15.9	317	16	US-10-795-927-6	Sequence 6, Appli
32	108.5	15.9	317	18	US-10-955-119-30	Sequence 30, Appli
33	106.5	15.6	368	16	US-10-856-888-7	Sequence 7, Appli
34	102	15.0	322	9	US-09-728-479-11	Sequence 11, Appli
35	102	15.0	322	18	US-10-958-169-17	Sequence 17, Appli
36	102	15.0	353	15	US-10-633-035-8	Sequence 8, Appli
37	102	15.0	353	18	US-10-958-169-18	Sequence 18, Appli
38	96.5	14.1	358	16	US-10-856-888-6	Sequence 6, Appli
39	96.5	14.1	358	16	US-10-856-888-16	Sequence 16, Appli
40	96.5	14.1	359	14	US-10-376-133-20	Sequence 20, Appli
41	96.5	14.1	359	18	US-10-712-124-92	Sequence 92, Appli
42	95.5	14.0	125	9	US-09-768-826-36	Sequence 36, Appli
43	95.5	14.0	125	11	US-09-833-245-2256	Sequence 256, Ap
44	95.5	14.0	125	16	US-10-874-484-36	Sequence 36, Appli
45	95.5	14.0	135	16	US-10-856-888-20	Sequence 20, Appli
46	95.5	14.0	136	9	US-09-728-479-9	Sequence 9, Appli
47	95.5	14.0	136	9	US-09-154-750A-79	Sequence 79, Appli
48	95.5	14.0	136	9	US-09-263-689-13	Sequence 13, Appli
49	95.5	14.0	136	14	US-10-235-674-13	Sequence 13, Appli
50	95.5	14.0	136	14	US-10-316-253-260	Sequence 260, App
51	95.5	14.0	136	15	US-10-133-234A-5	Sequence 2, Appli
52	95.5	14.0	136	15	US-10-398-519-15	Sequence 15, Appli
53	94.5	13.9	131	10	US-09-948-227-7	Sequence 7, Appli
54	94.5	13.9	131	16	US-10-795-927-7	Sequence 7, Appli
55	94	13.8	315	16	US-10-856-888-17	Sequence 17, Appli
56	93	13.6	262	15	US-10-133-234A-5	Sequence 5, Appli
57	91.5	13.4	316	16	US-10-856-888-9	Sequence 9, Appli
58	88.5	13.0	264	16	US-10-473-127-1448	Sequence 1448, Ap
59	87.5	12.8	300	15	US-10-398-519-17	Sequence 15, Appli
60	87.5	12.8	316	9	US-09-263-689-15	Sequence 17, Appli
61	87.5	12.8	316	9	US-09-263-689-17	Sequence 15, Appli
62	87.5	12.8	316	14	US-10-235-674-15	Sequence 17, Appli
63	87.5	12.8	316	14	US-10-235-674-17	Sequence 8, Appli
64	87.5	12.8	316	18	US-10-505-769-8	Sequence 8, Appli
65	87	12.8	140	18	US-10-450-763-53319	Sequence 53319, A
66	86.5	12.7	333	15	US-10-138-588-76	Sequence 76, Appli
67	86.5	12.7	333	17	US-10-506-047-2	Sequence 2, Appli
68	85.5	12.5	263	15	US-10-398-519-14	Sequence 14, Appli
69	84.5	12.4	315	9	US-09-728-479-10	Sequence 10, Appli
70	84.5	12.4	316	15	US-10-151-166-2	Sequence 2, Appli
71	84	12.3	241	9	US-09-738-626-6010	Sequence 6010, Ap
72	84	12.3	324	9	US-09-728-479-7	Sequence 7, Appli
73	84	12.3	324	9	US-09-263-689-11	Sequence 11, Appli
74	84	12.3	324	14	US-10-235-674-11	Sequence 11, Appli
75	82.5	12.1	136	15	US-10-133-234A-7	Sequence 7, Appli
76	82.5	12.1	136	15	US-10-133-234A-8	Sequence 8, Appli
77	82.5	12.1	250	16	US-10-133-127-1350	Sequence 1350, Ap
78	82.5	12.1	281	14	US-10-362-017-5	Sequence 5, Appli
79	81.5	12.0	145	14	US-10-362-017-4	Sequence 4, Appli
80	81.5	12.0	268	14	US-10-106-698-6399	Sequence 2, Appli
81	81.5	12.0	323	9	US-09-802-674-2	Sequence 2, Appli
82	81.5	12.0	323	9	US-09-922-217-1064	Sequence 1064, Ap
83	81.5	12.0	323	9	US-09-988-252-16	Sequence 16, Appli
84	81.5	12.0	323	9	US-09-833-263-1064	Sequence 1064, Ap

85 81.5 12.0 323 9 US-09-263-689-2 Sequence 2, Appli  
86 81.5 12.0 323 9 US-09-981-353-110 Sequence 110, App  
87 81.5 12.0 323 13 US-10-025-380-1064 Sequence 1064, Ap  
88 81.5 12.0 323 14 US-10-235-674-2 Sequence 2, Appli  
89 81.5 12.0 323 14 US-10-235-994-30 Sequence 30, Appl  
90 81.5 12.0 323 14 US-10-376-133-17 Sequence 17, Appl  
91 81.5 12.0 323 16 US-10-776-601-16 Sequence 16, Appl  
92 81.5 12.0 323 16 US-10-856-888-21 Sequence 21, Appl  
93 81.5 12.0 323 20 US-11-032-421-2 Sequence 2, Appli  
94 81.5 12.0 329 9 US-09-802-674-13 Sequence 13, Appl  
95 81.5 12.0 329 20 US-11-032-421-13 Sequence 13, Appl  
96 81.5 12.0 351 14 US-10-106-698-4502 Sequence 4502, Ap  
97 81 11.9 171 14 US-10-190-866A-2 Sequence 2, Appli  
98 81 11.9 401 14 US-10-190-866A-1 Sequence 1, Appli  
99 80.5 11.8 242 15 US-10-133-234A-4 Sequence 4, Appli  
100 80 11.7 145 9 US-09-894-526-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-10-156-136-29  
; Sequence 29, Application US/10156136  
; Publication No. US20030129696A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni, et al.  
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoding Receptors

; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Ave  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/156,136  
; FILING DATE: 29-May-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/010,146  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: WO US98/00959  
; FILING DATE: 21-JAN-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PF354PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Query Match 98.4%; Score 671; DB 14; Length 133;  
Best Local Similarity 99.2%; Pred. No. 6.1e-68;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
US-10-156-136-29  
1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFADRTL 60

Db 1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFADRTL 60  
Qy 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKALNGOGLGATSMNQALEQLREL 120  
Db 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKALNGOGLGATSMNQALEQLREL 120  
Qy 121 RISGSVQLYCVHS 133  
Db 121 RISGSVQLYCVHS 133

RESULT 2  
US-10-455-366-2  
; Sequence 2, Application US/10455366  
; Publication No. US20030208044A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Galectin 11  
; FILE REFERENCE: PF354P2  
; CURRENT APPLICATION NUMBER: US/10/455,366  
; CURRENT FILING DATE: 2003-06-06  
; PRIOR APPLICATION NUMBER: US/09/557,170  
; PRIOR FILING DATE: 2000-04-21  
; PRIOR APPLICATION NUMBER: 09/109,864  
; PRIOR FILING DATE: 1998-06-06  
; PRIOR APPLICATION NUMBER: 09/010,146  
; PRIOR FILING DATE: 1998-01-21  
; PRIOR APPLICATION NUMBER: 60/034,205  
; PRIOR FILING DATE: 1997-01-21  
; PRIOR APPLICATION NUMBER: 60/034,204  
; PRIOR FILING DATE: 1997-01-21  
; PRIOR APPLICATION NUMBER: 60/169,932  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR APPLICATION NUMBER: 60/130,390  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-455-366-2

Query Match 98.4%; Score 671; DB 15; Length 133;  
Best Local Similarity 99.2%; Pred. No. 6.1e-68;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
US-10-455-366-2

Qy 1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFADRTL 60  
Db 1 MSPRLEVPCHALPQGLSPGVIIVRGLVLPQPKHFTVSLRDQAAHAPVTLRASFADRTL 60  
Qy 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKALNGOGLGATSMNQALEQLREL 120  
Db 61 AMISRWGQKLLISAPPLFYPPQPFVLLLFQSGGLKALNGOGLGATSMNQALEQLREL 120  
Qy 121 RISGSVQLYCVHS 133  
Db 121 RISGSVQLYCVHS 133

RESULT 3  
US-11-041-419-29  
; Sequence 29, Application US/11041419  
; Publication No. US20050176044A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni, et al.  
; TITLE OF INVENTION: Polynucleotides And Polypeptides Encoding Receptors  
; FILE REFERENCE: PF354C3  
; CURRENT APPLICATION NUMBER: US/11/041,419  
; CURRENT FILING DATE: 2005-01-25  
; NUMBER OF SEQ ID NOS: 63  
; SEQ ID NO 29

```

; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-041-419-29

Query Match      98.4%; Score 671; DB 20; Length 133;
Best Local Similarity 99.2%; Pred. No. 6.1e-68;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||
Db      1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||

Qy      61 AWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGKLALNGQGIGATSMNQQALEQLREL 120
        |||||||
Db      61 AWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGKLALNGQGIGATSMNQQALEQLREL 120
        |||||||

Qy      121 RISGSVQLYCVHS 133
        |||||||
Db      121 RISGSVQLYCVHS 133
        |||||||

RESULT 4
US-10-455-366-25
; Sequence 25, Application US/10455366
; Publication No. US20030208044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/10/455,366
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/09/557,170
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-366-25

Query Match      98.4%; Score 671; DB 15; Length 275;
Best Local Similarity 99.2%; Pred. No. 1.5e-67;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||
Db      143 MSPRLEVPCHALPQGLSPGQVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 202
        |||||||

Qy      61 AWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGKLALNGQGIGATSMNQQALEQLREL 120
        |||||||
Db      203 AWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGKLALNGQGIGATSMNQQALEQLREL 262
        |||||||

Qy      121 RISGSVQLYCVHS 133
        |||||||
Db      263 RISGSVQLYCVHS 275
        |||||||

RESULT 5
US-10-455-366-27
```

```

; Sequence 27, Application US/10455366
; Publication No. US20030208044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/10/455,366
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/09/557,170
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/130,390
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-366-27

Query Match      98.4%; Score 671; DB 15; Length 296;
Best Local Similarity 99.2%; Pred. No. 1.7e-67;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MSPRLEVPCHALPQGLSPGQVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 60
        |||||||
Db      164 MSPRLEVPCHALPQGLSPGQVIIVRGLVLOEPKHPTVSLRDQAAHAPVTLRASFAADRTL 223
        |||||||

Qy      61 AWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGKLALNGQGIGATSMNQQALEQLREL 120
        |||||||
Db      224 AWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGKLALNGQGIGATSMNQQALEQLREL 283
        |||||||

Qy      121 RISGSVQLYCVHS 133
        |||||||
Db      284 RISGSVQLYCVHS 296
        |||||||

RESULT 6
US-09-747-804-1
; Sequence 1, Application US/09747804
; Patent No. US20010010913A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
; FILE REFERENCE: PF-0576 US
; CURRENT APPLICATION NUMBER: US/09/747,804
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/131,648
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2635136
US-09-747-804-1
```





```
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;
QY 3 PRLEVPCHALPQGLSPGQVIVIRGLVLPQPKPTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGDIAPHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYPPORFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLR 118
Db 71 TQINNSWGPEERSLPGSMPPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 13
US-10-235-674-12
; Sequence 12, Application US/10235674
; Publication No. US20030040081A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 9 and 10SV Polynucleotides
; FILE REFERENCE: 1488.0560004
; CURRENT APPLICATION NUMBER: US/10/235,674
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 09/656,450
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/263,689
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 08/946,914
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: US 60/028,093
; PRIOR FILING DATE: 1996-10-09
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Rat
US-10-235-674-12

Query Match 17.4%; Score 119; DB 14; Length 145;
Best Local Similarity 26.9%; Pred. No. 3.6e-05;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;
QY 3 PRLEVPCHALPQGLSPGQVIVIRGLVLPQPKPTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGDIAPHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYPPORFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLR 118
Db 71 TQINNSWGPEERSLPGSMPPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 14
US-10-156-136-42
; Sequence 42, Application US/10156136
; Publication No. US20030129696A1
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoding Receptors
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave
; CITY: Rockville
```

```
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/156,136
; APPLICATION NUMBER: US/09/010,146
; FILING DATE: <Unknown>
; APPLICATION NUMBER: WO US98/00959
; FILING DATE: 21-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF354PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-156-136-42

Query Match 17.4%; Score 119; DB 14; Length 145;
Best Local Similarity 26.9%; Pred. No. 3.6e-05;
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;
QY 3 PRLEVPCHALPQGLSPGQVIVIRGLVLPQPKPTVSLR---DQAAHAPVTLRASPADRT 59
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGDIAPHLNPRFDENAVVRN 70
QY 60 LAWISRWG-QKKLISAPFLFYPPORFFVLLLFQEGGLKALNGQGLGATSMNQQALEQLR 118
Db 71 TQINNSWGPEERSLPGSMPPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130
QY 119 ELRISGSVOL 128
Db 131 TLEVAGDIQL 140

RESULT 15
US-10-455-366-3
; Sequence 3, Application US/10455366
; Publication No. US20030208044A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Galectin 11
; FILE REFERENCE: PF354P2
; CURRENT APPLICATION NUMBER: US/10/455,366
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/09/557,170
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 09/109,864
; PRIOR FILING DATE: 1998-06-06
; PRIOR APPLICATION NUMBER: 09/010,146
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: 60/034,205
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/034,204
; PRIOR FILING DATE: 1997-01-21
; PRIOR APPLICATION NUMBER: 60/169,932
; PRIOR FILING DATE: 1999-12-10
```



;; PRIOR APPLICATION NUMBER: 60/130,390  
;; PRIOR FILING DATE: 1999-04-21  
;; NUMBER OF SEQ ID NOS: 27  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 3  
;; LENGTH: 145  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-455-366-3

Query Match 17.4%; Score 119; DB 15; Length 145;  
Best Local Similarity 26.9%; Pred. No. 3.6e-05;  
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;  
  
Qy 3 PRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHPVTSLR---DOAAHAPVTLRASFPADRT 59  
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGGIAFHLPNPFDENAVVRN 70  
  
Qy 60 LAWISRWG-QKKLISAPFLFYQRFVLLLFQGGKIALNGGGLGATSMNQQAELQLR 118  
Db 71 TQINNSWGPEERSLPGSMPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130  
  
Qy 119 ELRISGSVOL 128  
Db 131 TLEVAGDIQL 140

RESULT 16  
US-11-041-419-42  
;; Sequence 42, Application US/11041419  
;; Publication No. US20050176044A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ni et al.  
;; TITLE OF INVENTION: Polynucleotides And Polypeptides Encoding Receptors  
;; FILE REFERENCE: PF354C3  
;; CURRENT APPLICATION NUMBER: US/11/041.419  
;; CURRENT FILING DATE: 2005-01-25  
;; NUMBER OF SEQ ID NOS: 63  
;; SEQ ID NO 42  
;; LENGTH: 145  
;; TYPE: PRT  
;; ORGANISM: Rattus sp.  
US-11-041-419-42

Query Match 17.4%; Score 119; DB 20; Length 145;  
Best Local Similarity 26.9%; Pred. No. 3.6e-05;  
Matches 35; Conservative 26; Mismatches 65; Indels 4; Gaps 2;  
  
Qy 3 PRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHPVTSLR---DOAAHAPVTLRASFPADRT 59  
Db 11 PNLAVPFTSIPNGLYPSKSIIVSGVLSDAKRFQINLRCCGGIAFHLPNPFDENAVVRN 70  
  
Qy 60 LAWISRWG-QKKLISAPFLFYQRFVLLLFQGGKIALNGGGLGATSMNQQAELQLR 118  
Db 71 TQINNSWGPEERSLPGSMPSRGQRFVSWILCEGHCFKVAVDGQHICEYSHRLMNLDPIN 130  
  
Qy 119 ELRISGSVOL 128  
Db 131 TLEVAGDIQL 140

RESULT 17  
US-10-491-213-9  
;; Sequence 9, Application US/10491213  
;; Publication No. US20050048490A1  
;; GENERAL INFORMATION:  
;; APPLICANT: INCYTE CORPORATION; AZIMZAI, Yalda;  
;; APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;  
;; APPLICANT: BOWOMSKY, Mark L.; CHAWLA, Narinder K.;  
;; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;  
;; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.;  
;; APPLICANT: GORVAD, Ann E.; GRIFFIN, Jennifer A.;  
;; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;

;; APPLICANT: KABLE, Amy E.; KALAFUS, Daniel P.;  
;; APPLICANT: LEHR-MASON, Patricia M.; LU, Dyung Aina M.;  
;; APPLICANT: MARQUIS, Joseph P.; NGUYEN, Dannel B.;  
;; APPLICANT: RAMKUMAR, Jayalaxmi; RICHARDSON, Thomas W.;  
;; APPLICANT: KAREHT, Stephanie K.; SWARNAKAR, Anita;  
;; APPLICANT: TANG, Y. Tom; TRAN, Uyen K.;  
;; APPLICANT: WARREN, Bridget A.; XU, Yuming;  
;; APPLICANT: YAO, Monique G.; YUE, Huibin;  
;; APPLICANT: YUE, Henry  
;; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH  
;; FILE REFERENCE: PF-1213 USN  
;; CURRENT APPLICATION NUMBER: US/10/491,213  
;; CURRENT FILING DATE: 2004-03-26  
;; PRIOR APPLICATION NUMBER: PCT/US02/31095  
;; PRIOR FILING DATE: 2002-09-26  
;; PRIOR APPLICATION NUMBER: US 60/326,389  
;; PRIOR FILING DATE: 2001-09-28  
;; PRIOR APPLICATION NUMBER: US 60/327,380  
;; PRIOR FILING DATE: 2001-10-05  
;; PRIOR APPLICATION NUMBER: US 60/328,186  
;; PRIOR FILING DATE: 2001-10-05  
;; PRIOR APPLICATION NUMBER: US 60/329,690  
;; PRIOR FILING DATE: 2001-10-12  
;; PRIOR APPLICATION NUMBER: US 60/345,384  
;; PRIOR FILING DATE: 2001-10-26  
;; PRIOR APPLICATION NUMBER: US 60/348,165  
;; PRIOR FILING DATE: 2001-10-26  
;; PRIOR APPLICATION NUMBER: US 60/350,219  
;; PRIOR FILING DATE: 2001-11-02  
;; PRIOR APPLICATION NUMBER: US 60/344,518  
;; PRIOR FILING DATE: 2001-11-09  
;; PRIOR APPLICATION NUMBER: US 60/345,143  
;; PRIOR FILING DATE: 2001-11-09  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 114  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 9  
;; LENGTH: 158  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; OTHER INFORMATION: incyte ID No: 7500513CDI  
US-10-491-213-9  
  
Query Match 16.1%; Score 109.5; DB 17; Length 158;  
Best Local Similarity 25.5%; Pred. No. 0.00048;  
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;  
  
Qy 2 SPRLEVPCHALPQGLSPGVIIIVRGLVLOEPKHPVTSL-----RDOAAHAPVTLRASFA 56  
Db 21 TPQLSLPFAARLNTPMGPGRTVVVKGVEVNAKSFNVDLLAGKSKDIALHLNPLNLIKAF 80  
  
Qy 57 DRTLAWISRWGQKKLISAPFLFYQRFVLLLFQGGKIALNGGGLGATSMNQQAELQ 116  
Db 81 VRNSFLOESGEEERNITSPFSPGMFYEMIIYCDVREFKVAVNGVHSLEYKHFRLKLS 140  
  
Qy 117 LRELISGSVOLYCVHS 133  
Db 141 IDTLEINGDIHLEVR 157  
  
RESULT 18  
US-10-856-888-5  
;; Sequence 5, Application US/10856888  
;; Publication No. US20040235037A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Blumenfeld, Marta  
;; APPLICANT: Bougueleret, Lydie  
;; APPLICANT: Chumakov, Ilya  
;; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)  
;; FILE REFERENCE: GEN-1112XC1  
;; CURRENT APPLICATION NUMBER: US/10/856,888

```

; CURRENT FILING DATE: 2004-05-27
; PRIOR APPLICATION NUMBER: US/09/326,402
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 16
; OTHER INFORMATION: 5-2-162 : polymorphic amino acid Tyr or Phe
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 35
; OTHER INFORMATION: 5-2-213 : polymorphic amino acid Cys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 55
; OTHER INFORMATION: 5-3-84 : polymorphic amino acid Val or Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 163
; OTHER INFORMATION: 5-7-195 : polymorphic amino acid Ser or Arg
; US-10-856-888-5

Query Match      16.1%; Score 109.5; DB 16; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGOVLIIRGLVLQEPKHPVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGSKDIALHLNPLNIKAF 238

QY      57 DRTLAWISWGOKKLISAPFLFPQRFVLLFQEGGLKIALNGQGLGATSMNQALEQ 116
Db      239 VRNSFLOESWGEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLKYKHFKELESS 298

QY      117 LRELRISSGVOLYCVHS 133
Db      299 IDTLEINGDIHLLVRS 315

RESULT 19
US-10-856-888-15
; Sequence 15, Application US/10856888
; Publication No. US20040235037A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-112XC1
; CURRENT APPLICATION NUMBER: US/10/856,888
; PRIOR FILING DATE: 2004-05-27
; PRIOR APPLICATION NUMBER: US/09/326,402
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

Query Match      16.1%; Score 109.5; DB 16; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGOVLIIRGLVLQEPKHPVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGSKDIALHLNPLNIKAF 238

QY      57 DRTLAWISWGOKKLISAPFLFPQRFVLLFQEGGLKIALNGQGLGATSMNQALEQ 116
Db      239 VRNSFLOESWGEERNITSPFSPGMYFEMIIYCDVREFKAVNGVHSLKYKHFKELESS 298

QY      117 LRELRISSGVOLYCVHS 133
Db      299 IDTLEINGDIHLLVRS 315

RESULT 20
US-09-263-689-6
; Sequence 6, Application US/09263689
; Patent No. US20020150970A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Galectin 8, 9, 10 and 10SV
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein, & Fox P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,689
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/946,914
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0560001/EKS/SGW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-263-689-6

Query Match      16.1%; Score 109.5; DB 9; Length 317;
Best Local Similarity 25.5%; Pred. No. 0.0011;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

QY      2 SPRLEVPCHALPQGLSPGOVLIIRGLVLQEPKHPVSL-----RDQAAHAPVTLRASFA 56
Db      180 TPQLSLPFAARLNTPMGPGRTVVVKGEVNAKSFNVDLLAGSKDIALHLNPLNIKAF 239
```





```

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
; FILE REFERENCE: PP-0576 US
; CURRENT APPLICATION NUMBER: US/09/747,804
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/131,648
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g1932712
US-09-747-804-5

Query Match      15.9%; Score 108.5; DB 9; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0015;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy      2 SPRLEVPCHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVQGEVNAKSFNVDDLAKGSKDIALHLNPRLNKAF 238

Qy      57 DRTLAWISRMGQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db      239 VRNSFLOESWGEERNITSPFPSPGMYFEMIYCDVREFKVAVNGVHSLKYKHFKEKLS 298

Qy      117 LRELRISSVOLYCVHS 133
Db      299 IDTLEINGDIHLLEVR 315

RESULT 28
US-10-177-293-354
; Sequence 354, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002

```

```

; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-354

Query Match      15.9%; Score 108.5; DB 14; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0015;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy      2 SPRLEVPCHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVQGEVNAKSFNVDDLAKGSKDIALHLNPRLNKAF 238

Qy      57 DRTLAWISRMGQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db      239 VRNSFLOESWGEERNITSPFPSPGMYFEMIYCDVREFKVAVNGVHSLKYKHFKEKLS 298

Qy      117 LRELRISSVOLYCVHS 133
Db      299 IDTLEINGDIHLLEVR 315

RESULT 29
US-10-482-029-134
; Sequence 134, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482,029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-482-029-134

Query Match      15.9%; Score 108.5; DB 17; Length 316;
Best Local Similarity 25.5%; Pred. No. 0.0015;
Matches 35; Conservative 23; Mismatches 74; Indels 5; Gaps 1;

Qy      2 SPRLEVPCHALPQGLSPGQVIVIRGLVLOEPKHPTVSL-----RDQAAHAPVTLRASFA 56
Db      179 TPQLSLPFAARLNTPMGPGRTVVQGEVNAKSFNVDDLAKGSKDIALHLNPRLNKAF 238

Qy      57 DRTLAWISRMGQKKLISAPFLFYQRFVEVLLLFQEGGLKALNGOGLGATSMNQALEQ 116
Db      239 VRNSFLOESWGEERNITSPFPSPGMYFEMIYCDVREFKVAVNGVHSLKYKHFKEKLS 298

Qy      117 LRELRISSVOLYCVHS 133
Db      299 IDTLEINGDIHLLEVR 315

RESULT 30
US-09-948-227-6
; Sequence 6, Application US/09948227
; Publication No. US20030050266A1
; GENERAL INFORMATION:
; APPLICANT: Paul B. Fisher
; APPLICANT: Raul V. Gopalakrishnan
; TITLE OF INVENTION: Anti-tumor Effects of Prostate Carcinoma

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:32:58 ; Search time 16 Seconds  
(without alignments)  
799.802 Million cell updates/sec

Title: HJACE54  
Perfect score: 682  
Sequence: 1 msprlvpcshlpqglspg.....leqlrelrsgsvqlcvs 133

Scoring table: BLOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	17.4	145	2	galectin-5 - rat
2	108.5	15.9	317	2	prostate carcinoma
3	95.5	14.0	136	2	galectin-7 - human
4	94.5	13.9	727	2	heterocyst-specifi
5	94.5	13.9	727	2	heterocyst-specifi
6	87.5	12.8	316	2	galectin-8 - rat
7	85.5	12.5	264	2	galactose-specific
8	85.5	12.5	264	2	galactose-specific
9	84	12.3	324	2	lactose-binding le
10	81	11.9	848	2	interleukin-4-indu
11	80.5	11.8	242	2	galectin-3 - rabbi
12	78.5	11.5	262	2	IGF-binding protei
13	78.5	11.5	323	2	lectin L-36 - pig
14	77.5	11.4	245	2	carbohydrate-bind
15	77.5	11.4	250	2	galectin 3 - human
16	76.5	11.2	294	2	lactose-binding le
17	74.5	10.9	129	2	ig kappa chain - h
18	72.5	10.6	191	2	hypothetical prote
19	72.5	10.6	279	2	beta-galactoside-b
20	72.5	10.6	285	2	hypothetical prote
21	72.5	10.6	316	1	nonstructural prot
22	72.5	10.6	645	2	asparagine synthas
23	72	10.6	131	2	T cell receptor Ck
24	72	10.6	135	1	LNBOGB
25	72	10.6	361	2	AG0052
26	71	10.4	135	1	LNHUGB
27	71	10.4	288	2	C75570
28	71	10.4	392	2	F72632
29	71	10.4	396	2	B83188

30	71	10.4	468	2	AI2811	serine proteinase
31	71	10.4	495	2	B97590	htra protein homol
32	70.5	10.3	688	2	H96681	protein FIE2.10 l
33	70	10.3	303	2	D95011	hypothetical prote
34	70	10.3	320	2	G97882	hypothetical prote
35	69.5	10.2	746	2	T35811	probable phosphodi
36	69	10.1	119	2	G32516	T-cell receptor al
37	69	10.1	165	2	AH2010	molybdopterin conv
38	69	10.1	240	2	AD2534	hypothetical prote
39	69	10.1	285	2	F88281	protein ZK892.1 l
40	69	10.1	297	2	T28073	hypothetical prote
41	69	10.1	310	2	T33972	hypothetical prote
42	69	10.1	544	2	A95591	hypothetical prote
43	68.5	10.0	236	2	AE0167	probable leucyl/ph
44	68.5	10.0	302	2	D82996	probable transcrip
45	68	10.0	323	2	T26796	hypothetical prote
46	68	10.0	409	1	S75020	acetyl-CoA C-acety
47	68	10.0	457	2	T05651	hypothetical prote
48	68	10.0	588	2	A49618	probable ataxia-te
49	67.5	9.9	203	2	D81116	RNA methyltransfer
50	67.5	9.9	261	2	B81904	probable RNA methy
51	67.5	9.9	398	2	AC3578	glycine betaine/l-
52	67.5	9.9	461	1	A46077	steroid hormone re
53	67.5	9.9	554	2	S46346	gag polyprotein -
54	67.5	9.9	790	1	PLPG	plasmin (EC 3.4.21
55	67.5	9.9	1688	1	C69224	cobalamin biosynth
56	67	9.8	244	2	AG1013	probable exported
57	67	9.8	356	2	AG2907	conserved hypother
58	67	9.8	378	2	G97682	hypothetical prote
59	67	9.8	488	2	G65216	hypothetical 53.4
60	67	9.8	489	2	AG3038	conserved hypother
61	67	9.8	515	2	D98247	mexT protein (AP11
62	67	9.8	946	2	S27921	nuclear antigen EB
63	67	9.8	1123	1	WMBEH7	UL37 protein - hum
64	67	9.8	1339	2	T40245	probable transcrip
65	66.5	9.8	369	2	T30592	conserved hypother
66	66.5	9.8	666	2	G82130	2,4-dienoyl-CoA re
67	66.5	9.8	746	2	S34656	amine oxidase (cop
68	66	9.7	219	2	AF0639	flagellar basal bo
69	66	9.7	349	2	AF2130	iron(III) dicitrat
70	66	9.7	398	2	D83956	hypothetical prote
71	66	9.7	624	2	H71316	probable penicilli
72	66	9.7	642	2	S58154	hypothetical prote
73	66	9.7	696	2	S61502	flagellar biosynth
74	66	9.7	703	2	B41538	ABC-type transport
75	65.5	9.6	159	2	I84615	coagulation factor
76	65.5	9.6	283	2	T19933	hypothetical prote
77	65.5	9.6	392	2	S56557	hypothetical 41.4K
78	65.5	9.6	395	2	F83873	hypothetical prote
79	65.5	9.6	441	2	AH0603	conserved hypother
80	65.5	9.6	646	2	AD0912	probable lipoprote
81	65.5	9.6	694	2	A95866	probable transketo
82	65.5	9.6	893	2	AG3022	two component sens
83	65.5	9.6	900	2	C98262	polyketide synthas
84	65.5	9.6	2124	2	T28658	beta-galactoside-b
85	65	9.5	135	1	LNRTGB	NAD(P)-arginine AD
86	65	9.5	300	2	JC4367	hypothetical prote
87	65	9.5	303	2	T29513	probable histidino
88	65	9.5	384	2	E72507	3-deoxy-D-manno-oc
89	65	9.5	425	2	AI0971	hypothetical prote
90	65	9.5	437	2	AH2159	abdominal segment
91	65	9.5	1533	2	A46221	SA85-1.2 protein -
92	64.5	9.5	240	2	S11293	hypothetical prote
93	64.5	9.5	284	2	A70976	group 2 sigma 70-t
94	64.5	9.5	316	2	AC2500	probable transport
95	64.5	9.5	392	2	E91290	probable transport
96	64.5	9.5	392	2	H86131	coagulation factor
97	64.5	9.5	466	1	KPHU7	acetolactate synth
98	64.5	9.5	585	2	C70341	ribonuclease HII V
99	64	9.4	206	2	H82100	ribulose-phosphate
100	64	9.4	225	2	AG1000	





Db 242 SFLODANGEERNITCFPPSSGMYFEMIICYDVRFEKVAVNGVHSLEYKHKRFDLSSIDT 301

Qy 120 LRISGSVOLYCVHS 133  
 Db 302 LAVDGDRIILLDVR 315

RESULT 7

A28651  
 galactose-specific lectin - mouse  
 N;Alternate names: carbohydrate-binding protein 35; IgE-binding protein; lectin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 28-Aug-1989 #sequence revision 10-Feb-1995 #text\_change 09-Jul-2004  
 C;Accession: S08537; A28651; A37385; A35185  
 R;Cherayil, B.J.; Weiner, S.J.; Pillai, S.  
 J. Exp. Med. 170, 1959-1972, 1989  
 A;Title: The Mac-2 antigen is a galactose-specific lectin that binds IgE.  
 A;Reference number: S08537; MUID:90063462; PMID:2584931  
 A;Accession: S08537  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-264 <R>  
 A;Cross-references: UNIPROT:P16110; EMBL:X16834; NID:G52986; PIDN:CAA34736.1; J  
 A;Note: this sequence was submitted to the EMBL Data Library, Oct-1989  
 R;Jia, S.; Wang, J.L.  
 J. Biol. Chem. 263, 6009-6011, 1988  
 A;Title: Carbohydrate binding protein 35. Complementary DNA sequence reveals ho  
 A;Reference number: A28651; MUID:88198129; PMID:3360772  
 A;Accession: A28651  
 A;Molecule type: mRNA  
 A;Residues: 1-3-264 <JIA>  
 A;Cross-references: EMBL:J03723  
 R;Raz, A.; Pazerini, G.; Carmi, P.  
 Cancer Res. 49, 3489-3493, 1989  
 A;Title: Identification of the metastasis-associated, galactoside-binding lecti  
 A;Reference number: A37385; MUID:89275058; PMID:2525069  
 A;Accession: A37385  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-3,'T',5-91,'ST',94-109,'SAP',113-264 <RAZ>  
 A;Cross-references: GB:X16074; NID:G52850; PIDN:CAA34206.1; PID:G52851  
 A;Note: authors translated the codon GAA for residue 219 as Ala, GAC for residu  
 R;Woo, H.J.; Shaw, L.M.; Messier, J.M.; Mercurio, A.M.  
 J. Biol. Chem. 265, 7097-7099, 1990  
 A;Title: The major non-integrin laminin binding protein of macrophages is ident  
 A;Reference number: A35185; MUID:90236991; PMID:2332426  
 A;Accession: A35185  
 A;Molecule type: protein  
 A;Residues: 159-163;166-175;214-226 <WOO>  
 C;Superfamily: beta-galactoside-binding lectin  
 C;Keywords: lectin; phosphoprotein

Query Match 12.5%; Score 85.5; DB 2; Length 264;  
 Best Local Similarity 25.4%; Pred. No. 0.21;  
 Matches 34; Conservative 22; Mismatches 63; Indels 15; Gaps 4;

Qy 5 LEYCPSHALPQGLSPGVIIIVRGILVLOEPKHPTVSLR---DQAAHAPVTLRASFADRTLA 61  
 Db 128 LTPYDPLPGGVMPRLITIMGTGVKNANRVLDFRRGNDVAFH----FNPFRNENRR 183

Qy 62 WI-----SRGQKKLISAPFLFYPPORFEVLLLFQEGGLKALNGOGLGATSMNQAL 114  
 Db 184 VIVCNTRQDNNWGKEERQSA-FPFESGKPFKIQLVLEADHFKVAVNDAHLLQYNHRMKNL 242

Qy 115 EQLRELRISGSVOL 128  
 Db 243 REISQLGISGDITL 256

RESULT 8

A45983  
 lactose-binding lectin Mac-2 - mouse  
 C;Species: Mus musculus (house mouse)

C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: A45983  
R;Rosenberg, I.M.; Iyer, R.; Cherayil, B.; Chiodino, C.; Pillai, S.  
J. Biol. Chem. 268, 12393-12400, 1993  
A;Title: Structure of the murine Mac-2 gene. Splice variants encode proteins lacking functionally important regions.  
A;Reference number: A45983; MUID:93286070; PMID:8509379  
A;Accession: A45983  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-264 <ROS>  
A;Cross-references: UNIPROT:Q8C253; GB:L08649  
A;Introns: 6/3; 128/3; 158/2; 209/2; 213/3  
C;Superfamily: beta-galactoside-binding lectin

Query Match 12.5%; Score 85.5; DB 2; Length 264;  
Best Local Similarity 25.4%; Pred. No. 0.21;  
Matches 34; Conservative 22; Mismatches 63; Indels 15; Gaps 4;

QY 5 LEVPCSHALPQGLSPGVIIIRGLVLQEPKHPVSLR---DQAAHAPVTLRASFAADRTLA 61  
Db 128 LTPVYDLPPLPGGVMPRLITIMGTVPKNANRIVLDFRGNDAVFH----FNPRFNNRR 183  
QY 62 WI-----SRWGQKLLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQAL 114  
Db 184 VIVCNTKQDNWNGEERQSA-FFESGKPKIQVLVEADHFKVAVNDHALHLLQYNHRMKNL 242  
QY 115 EQLRELRISGSVOL 128  
Db 243 REISQLGISDITL 256

RESULT 9  
A46631  
lactose-binding lectin L-36 - rat  
N;Alternate names: galectin-4  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A46631; S69096  
R;Oda, Y.; Herrmann, J.; Gitt, M.A.; Turck, C.W.; Burlingame, A.L.; Barondes, S.H.; Lefebvre, J. Biol. Chem. 268, 5929-5939, 1993  
A;Title: Soluble lactose-binding lectin from rat intestine with two different carbohydrate specificities.  
A;Reference number: A46631; MUID:93194902; PMID:8449956  
A;Accession: A46631  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-324 <ODA>  
A;Cross-references: UNIPROT:P38552; GB:M73553; NID:g294571; PIDN:AAA41505.1; PID:g294572  
A;Experimental source: intestine  
A;Note: sequence extracted from NCBI backbone (NCBIN:128409, NCBIP:128410)  
R;Hardy, F.; Deviller, P.; Louisot, P.; Martin, A.  
FEBS Lett. 359, 169-172, 1995  
A;Title: Purification and characterization of the N-terminal domain of galectin-4 from rat intestine.  
A;Reference number: S69096; MUID:95172227; PMID:7867792  
A;Accession: S69096  
A;Molecule type: protein  
A;Residues: 13-37;44-50,'E',52-66 <TAR>  
C;Superfamily: lactose-binding lectin L-36  
C;Keywords: lectin

Query Match 12.3%; Score 84; DB 2; Length 324;  
Best Local Similarity 26.1%; Pred. No. 0.39;  
Matches 35; Conservative 23; Mismatches 56; Indels 20; Gaps 5;

QY 7 VPCSHALPQGLSPGVIIIRGLVLQEPKHPVSLR-----DQAAH-----APVTLRASFA 56  
Db 194 VPVVGTLQGLTARRTIKGYVLPATAKNLIINFKVSGTGDIAFHMNPRIGDCVWRNSYM 253  
QY 57 DRTLWISRWG--QKLLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQAL 114  
Db 254 N-----GSWGSERKIPYNP--FGAGQFFDLIRCTGDRFKVFGANGQHLFDPSHRFOAF 305  
QY 115 EQLRELRISGSVOL 128

C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: A45983  
R;Rosenberg, I.M.; Iyer, R.; Cherayil, B.; Chiodino, C.; Pillai, S.  
J. Biol. Chem. 268, 12393-12400, 1993  
A;Title: Structure of the murine Mac-2 gene. Splice variants encode proteins lacking functionally important regions.  
A;Reference number: A45983; MUID:93286070; PMID:8509379  
A;Accession: A45983  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-264 <ROS>  
A;Cross-references: UNIPROT:Q8C253; GB:L08649  
A;Introns: 6/3; 128/3; 158/2; 209/2; 213/3  
C;Superfamily: beta-galactoside-binding lectin

Query Match 12.5%; Score 85.5; DB 2; Length 264;  
Best Local Similarity 25.4%; Pred. No. 0.21;  
Matches 34; Conservative 22; Mismatches 63; Indels 15; Gaps 4;

QY 5 LEVPCSHALPQGLSPGVIIIRGLVLQEPKHPVSLR---DQAAHAPVTLRASFAADRTLA 61  
Db 128 LTPVYDLPPLPGGVMPRLITIMGTVPKNANRIVLDFRGNDAVFH----FNPRFNNRR 183  
QY 62 WI-----SRWGQKLLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQAL 114  
Db 184 VIVCNTKQDNWNGEERQSA-FFESGKPKIQVLVEADHFKVAVNDHALHLLQYNHRMKNL 242  
QY 115 EQLRELRISGSVOL 128  
Db 243 REISQLGISDITL 256

RESULT 9  
A46631  
lactose-binding lectin L-36 - rat  
N;Alternate names: galectin-4  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A46631; S69096  
R;Oda, Y.; Herrmann, J.; Gitt, M.A.; Turck, C.W.; Burlingame, A.L.; Barondes, S.H.; Lefebvre, J. Biol. Chem. 268, 5929-5939, 1993  
A;Title: Soluble lactose-binding lectin from rat intestine with two different carbohydrate specificities.  
A;Reference number: A46631; MUID:93194902; PMID:8449956  
A;Accession: A46631  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-324 <ODA>  
A;Cross-references: UNIPROT:P38552; GB:M73553; NID:g294571; PIDN:AAA41505.1; PID:g294572  
A;Experimental source: intestine  
A;Note: sequence extracted from NCBI backbone (NCBIN:128409, NCBIP:128410)  
R;Hardy, F.; Deviller, P.; Louisot, P.; Martin, A.  
FEBS Lett. 359, 169-172, 1995  
A;Title: Purification and characterization of the N-terminal domain of galectin-4 from rat intestine.  
A;Reference number: S69096; MUID:95172227; PMID:7867792  
A;Accession: S69096  
A;Molecule type: protein  
A;Residues: 13-37;44-50,'E',52-66 <TAR>  
C;Superfamily: lactose-binding lectin L-36  
C;Keywords: lectin

Query Match 12.3%; Score 84; DB 2; Length 324;  
Best Local Similarity 26.1%; Pred. No. 0.39;  
Matches 35; Conservative 23; Mismatches 56; Indels 20; Gaps 5;

QY 7 VPCSHALPQGLSPGVIIIRGLVLQEPKHPVSLR-----DQAAH-----APVTLRASFA 56  
Db 194 VPVVGTLQGLTARRTIKGYVLPATAKNLIINFKVSGTGDIAFHMNPRIGDCVWRNSYM 253  
QY 57 DRTLWISRWG--QKLLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQAL 114  
Db 254 N-----GSWGSERKIPYNP--FGAGQFFDLIRCTGDRFKVFGANGQHLFDPSHRFOAF 305  
QY 115 EQLRELRISGSVOL 128

Db 306 QRVDMLKIKDITL 319  
RESULT 10  
A54740  
interleukin-4-induced transcription factor stat - human  
C;Species: Homo sapiens (man)  
C;Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.  
Science 265, 1701-1706, 1994  
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.  
A;Reference number: A54740; MUID:94367369; PMID:8085155  
A;Accession: A54740  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-848 <HOU>  
A;Cross-references: UNIPROT:P42226  
C;Superfamily: human signal transducer and transcription activator STAT5A  
C;Keywords: DNA binding; transcription regulation

Query Match 11.9%; Score 81; DB 2; Length 848;  
Best Local Similarity 29.9%; Pred. No. 2.4;  
Matches 44; Conservative 14; Mismatches 49; Indels 40; Gaps 9;

QY 1 MSPRLVPCSHALP--QGLSPGVIIIRGLVLQEP-----KHPTVSL- 40  
Db 688 MYPQVYPHSHSIPPYQGLSPESVNVLS-AFQEPHLQMPPLSGQMSLFPDQHPQGLLP 746  
QY 41 ---RDQAAHAPVTLRASFAADRTLAWISRWGQKLLISAPFLFYQ-----RFPEVLLLFQ 91  
Db 747 CQPEHGVSSPDLCLS--DVTWV-----EBSCLSQVTAFFQGTWIGEDIFPPLLPPT 798  
QY 92 EGGL-KLALNGQ-LGATSMNQALEQ 116  
Db 799 EQDLTKLLEGQSGSGSLGAQPLLQ 825

RESULT 11  
galectin-3 - rabbit  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 16-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: JC4300  
R;Gaudin, J.C.; Monsigny, M.; Legrand, A.  
Gene 163, 249-252, 1995  
A;Title: Cloning of the cDNA encoding rabbit galectin-3.  
A;Reference number: JC4300; MUID:96011642; PMID:7590275  
A;Accession: JC4300  
A;Molecule type: mRNA  
A;Residues: 1-242 <GAU>  
A;Cross-references: UNIPROT:P47845; GB:U06470; NID:G606794; PIDN:AAC48491.1; PID:G606795  
A;Experimental source: vascular smooth muscle cells  
A;Note: The authors translated the codon TTC for residue 155 as Leu  
C;Comment: This protein has the functions on cell adhesion and proliferation. It is a su  
C;Genetics:  
A;Gene: Igals3  
C;Superfamily: beta-galactoside-binding lectin  
C;Keywords: muscle

Query Match 11.8%; Score 80.5; DB 2; Length 242;  
Best Local Similarity 25.0%; Pred. No. 0.65;  
Matches 34; Conservative 25; Mismatches 58; Indels 19; Gaps 5;

QY 5 LEVPCSHALPQGLSPGVIIIRGLVLQEPKHPVSL-----RDQAAHAPVTLRASFAADRT 59  
Db 106 LPVPYDLPPLPGGVMPRLITIVGTV--KENANRLALDFKRGNDVAFH----FNPRFNN 159  
QY 60 LAWI-----SRWGQKLLISAPFLFYQRFPEVLLLFQEGGLKALNGQGLGATSMNQ 112  
Db 160 RRIVCNTKVDNNWNGREER-QTTFPFEIGKPKIQVLVEADHFKVAVNDHALHLLQYNHRM 218

	Query Match	11.4%	Score 77.5	DB 2	Length 245
	Best Local Similarity	25.0%	Pred. NO. 1.4		
	Matches 34	Conservative 23	Mismatches 60	Indels 19	Gaps 5
Qy	5	LEVPCSHALPGLSGQGVIIIVRGLVLPQPKPTVSL-----RDQAAHAPVTLRASADRT	59		
Db	109	LTVPYKPLAGVMPRMILITMGTV--RPENARIILNLRGNDIAFH-----FNPRFENN	162		
Qy	60	LAWT-----SRWGQKKLISAPFLYFPQRFVEVILLFQEGGKLKALNGQGLGATSMNQ	112		
Db	163	RRVIVCNTKQDNNWGREQSA-PFFESGRPKIQVLVEADHPKVAVND AHLQYNHRMK	221		
Qy	113	ALEQLRELRIISGVQL	128		
Db	222	NLRINOMEISGDITL	237		

RESULT 15  
A35820  
galactin 3 - human  
N:Alternate names: CBP 35; epithelial-specific lectin 35; galactoside-binding lectin L-2  
C:Species: Homo sapiens (man)  
C;Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 09-Jul-2004  
R:Accession: A35820; JQ0916; A47473; A36071; A49800  
R:Robertson, M.W.; Albrandt, K.; Keller, D.; Liu, F.T.  
Biochemistry 29, 8093-8100, 1990  
A:Title: Human IGE-binding protein: a soluble lectin exhibiting a highly conserved inter  
A:Reference number: A35820; MUID:91084480; PMID:2261464  
A:Accession: A35820  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-250 <ROB>  
A:Cross-references: UNIPROT:P17931; GB:M57710; NID:G179530; PIDN:AAA35607.1; PID:G179531  
R:Oda, Y.; Leffler, H.; Sakakura, Y.; Kasai, K.; Barondes, S.H.  
Gene 99, 279-283, 1991  
A:Title: Human breast carcinoma cDNA encoding a galactoside-binding lectin homologous to  
A:Reference number: JQ0916; MUID:91216471; PMID:2022338

Gene 99, 219-263, 1991  
A>Title: Human breast carcinoma CDNA encoding a galactoside-binding lectin homologous to  
A.Reference number: JQ0916; MUID:91216471; PMID:2022338  
A.Accession: JQ0916  
A.Molecule type: mRNA  
A.Residues: 1-250 <CDA>  
A.Cross-references: GB:M36682; NID:g186921; PIDN:AAA36163.1; PID:g186922  
R.Lotz, M.M.; Andrews Jr., C.W.; Korzelius, C.A.; Lee, E.C.; Steele Jr., G.D.; Clarke, A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3466-3470, 1993  
A>Title: Decreased expression of Mac-2 (carbohydrate binding protein 35) and loss of its  
A.Reference number: A47473; MUID:93234518; PMID:7682704

A:Accession: A47473  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-63, 'P', 65-97, 'T', 99-250 <LOT>  
A:Cross-references: GB:S59012; NID:G299601; PIDN:AAB26229.1; PID:G299602  
A:Experimental source: normal colonic mucosa, colon carcinoma, cell line clone A  
A>Note: sequence extracted from NCBI backbone (NCBIN:129689, NCBIP:129692)  
R:Cherayil, B.J.; Chaitovitz, S.; Wong, C.; Pillai, S.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7324-7328, 1990  
A:Title: Molecular cloning of a human macrophage lectin specific for galactose.  
A:Reference number: A36071; MUID:90384999; PMID:2402511  
A:Accession: A36071  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-32, 'Q', 34, 'L', 37, 'RGFLSWGL', 46, 'RAGT', 51, 'R', 53-63, 'P', 65-87, 89-250 <CHB>  
A:Cross-references: GB:M35368; NID:g1196441  
A>Note: the sequence is revised in GenBank entry HUMMAC2, release 113.0, PIDN:AAA8086.1  
R:Raz, A.; Carmi, P.; Raz, T.; Hogan, V.; Mohamed, A.; Wolman, S.R.  
Cancer Res. 51, 2173-2178, 1991  
A:Title: Molecular cloning and chromosomal mapping of a human galactoside-binding protein  
A:Reference number: A49800; MUID:91183475; PMID:2009535  
A:Accession: A49800  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-32, 'Q', 34, 'LPGASVPGAYPAGTPGLSWTAPPGATMEHVELIRSTCTWSLRTQ', 86-104, 'A', 106, 'M'  
A:Cross-references: GB:M64303; NID:g413862  
A>Note: this translation is not annotated in GenBank entry HUMGALBIN, release 113.0  
C:Genetics:  
A:Gene: GDB:LGALS3; MAC-2; LGALS2  
A:Cross-references: GDB:127515; OMIM:137033  
A:Map position: lp13-lp13  
A:Superfamily: beta-galactoside-binding lectin  
C:Keywords: lectin; nucleus; phosphoprotein

Query Match 11.4%; Score 77.5; DB 2; Length 250;  
Best Local Similarity 25.0%; Pred. No. 1.4;  
Matches 34; Conservative 25; Mismatches 58; Indels 19; Gaps 5;

QY 5 LEVPCSHALPQGLSPQGVIVRGVLVQEPKHTVSL-----RQAAHAPVTLRASADRT 59  
DB 114 LIVPVNLPFGGVKPRMLITIGTV--KPNANRIALDFQRGNDVAFH-----FNPRFENN 167

QY 60 LAWI-----SRWGOKLISAPFLYPQRFVLLVFOGGKLALNGQGLGATSMNQ 112  
DB 168 RRIVVNTKLDNNNGREKQSV--FPFESGPKFIQVLVDPEDHKFAVNDHAHLLOYNHRVK 226

QY 113 ALEQLRELRIKSGVOL 128  
DB 227 KUNEISKLGISGIDIL 242

RESULT 16  
A49688  
lactose-binding lectin L-29 - dog  
C:Species: Canis lupus familiaris (dog)  
C>Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 24-Nov-1999  
C:Accession: A49688  
R:Herrmann, J.; Turck, C.W.; Atchison, R.E.; Huflejt, M.E.; Poulter, L.; Gitt, M.A.; Bur  
J. Biol. Chem. 268, 26704-26711, 1993  
A:Title: Primary structure of the soluble lactose binding lectin L-29 from rat and dog a  
agenses.  
A:Reference number: A49688; MUID:94075368; PMID:8253805  
A:Accession: A49688  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-294 <HER>  
A:Cross-references: GB:L23429  
C:Superfamily: beta-galactoside-binding lectin  
C:Keywords: acetylated amino end  
F:1/Modified site: acetylated amino end

Query Match 11.2%; Score 76.5; DB 2; Length 294;  
Best Local Similarity 25.0%; Pred. No. 2.2;

Matches 34; Conservative 25; Mismatches 58; Indels 19; Gaps 6;  
QY 5 LEVPCSHALPQGLSPQGVIVRGVLVQEPKHTVSL-----RQAAHAPVTLRASAD-- 57  
DB 158 LTVVPDLPFGGVKPRMLITIGTV--RPSANRLALDFKRGNDVAFH-----FNPRFENN 211  
QY 58 -RTLAWSR---WGOKLISAPFLYPQRFVLLVFOGGKLALNGQGLGATSMNQ 112  
DB 212 KRIVVNTKLDNNNGKEER-QAAPPFESGPKFIQVLVDPEDHKFAVNDHAHLLOYNHRMK 270  
QY 113 ALEQLRELRIKSGVOL 128  
DB 271 NLPETISKLGISGIDIL 286

RESULT 17  
IG kappa chain - human  
C:Species: Homo sapiens (man)  
C>Date: 06-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
C:Accession: S40332  
R:Klein, R.; Jaenichen, R.; Zachau, H.G.  
Eur. J. Immunol. 23, 3248-3271, 1993  
A:Title: Expressed human immunoglobulin chi genes and their hypermutation.  
A:Reference number: S40312; MUID:94080891; PMID:8258341  
A:Accession: S40332  
A>Status: preliminary; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-129 <KLE>  
A:Cross-references: EMBL:X72442; NID:g441352; PIDN:CAAS1110.1; PID:g441353  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:33-107/Domain: immunoglobulin homology <IMM>

Query Match 10.9%; Score 74.5; DB 2; Length 129;  
Best Local Similarity 29.9%; Pred. No. 1.4;  
Matches 23; Conservative 12; Mismatches 17; Indels 25; Gaps 4;

QY 43 QAAHAPVTLRASADRTL-----AWISRWGOKLISAPFLYPQRFVLLVFO-- 91  
DB 20 QMTQSPSTLSASVGDVRVIMTCRASIGSWLA-WYQOKLGKAP-----KLIIYKA 68

QY 92 ---EGGKLALNGQGLG 105  
DB 69 STLEGVPSRFSRGSG 85

RESULT 18  
S76352  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S76352  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S76352  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-191 <KAN>  
A:Cross-references: UNIPROT:Q55678; EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAAL0204;  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: Synechocystis hypothetical protein sl11956

Query Match 10.6%; Score 72.5; DB 2; Length 191;  
Best Local Similarity 28.7%; Pred. No. 3.5;  
Matches 31; Conservative 13; Mismatches 39; Indels 25; Gaps 6;

QY 3 PRLEVPCSHALPQGLSPGQVIVRGVLVQEP-----KHPTVSLR-----DQAAHA 47

asparagine synthase (glutamine-hydrolysing) (EC 6.3.5.4) - *Sinorhizobium meliloti* (strain C) [JGI]  
 C:Species: *Sinorhizobium meliloti*  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
 C:Accession: H96011  
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernar  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1.683-kb pSYMB megaplasmid from the N2-fixing endo  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: H96011  
 A:Status: preliminary  
 A:Molecule type: DNA

A;Residues: 1-645 <KUR>  
A;Cross-references: UNIPROT:Q92TX3; GB:AL591985; PIDN:CAC49760.1; PID:gl5141247; GSPDB:G  
A;Experimental source: strain 1021, megaplasmid pSYMB  
R;Galberv, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
C;Contents: annotation  
C;Genetics:  
A;Gene: aanB; SMB20652  
A;Genome: plasmid  
C;Keywords: ligase

Query Match 10.6%; Score 72.5; DB 2; Length 645;  
Best Local Similarity 21.7%; Pred. No. 14;  
Matches 30; Conservative 21; Mismatches 66; Indels 21; Gaps 4;

QY 4 RLEVPCHALPQGLSRQGV-IIVRGLVLQ-----EPKHTVSLRDOAA-----H 46  
DB 254 RADVPVGYALSGGLSDSIISALAGMTSQGLRTFSVTFDSAEDSAFOEEMAAALGTEH 313  
QY 47 APVTLRASPADRTLAMISRGQKLLISAPLPFPQRFVEVLLLFQGGGLKALNGQGLGA 106  
DB 314 RAVACRAGDIARFPDPIRTEPIRTA-----PAPLYKLSGLVRAGLKVLTGTGADE 369  
QY 107 TSMNQALQLELRISG 124  
DB 370 VPAGYDIFKEARVRFCG 387

RESULT 23  
S57871  
T cell receptor Ck4 alpha chain - human  
C;Species: Homo sapiens (man)  
C;Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 30-May-1997  
C;Accession: S57871  
R;Giegerich, G.; Pette, M.; Meinel, E.; Epplen, J.T.; Wekerle, H.; Hinkkanen, A.  
Eur. J. Immunol. 22, 753-758, 1992  
A;Title: Diversity of T cell receptor alpha and beta chain genes expressed by human T ce  
A;Reference number: S57869; MUID:92192091; PMID:1372258  
A;Accession: S57871  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-131 <GIE>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: T-cell receptor

Query Match 10.6%; Score 72; DB 2; Length 131;  
Best Local Similarity 33.9%; Pred. No. 2.6;  
Matches 39; Conservative 14; Mismatches 30; Indels 32; Gaps 10;

QY 16 GLSPGVIIVRGH-VLQEPKHPTVSLRDOAAHAPVTLRASPAD--RTLAWISR--WGQKK 70  
DB 10 GLLSAQVCCVRGIVQEQSP--PLLILQEGANS---TLRCNFSVNNLQWFHQNPWQ-- 62  
QY 71 LISAPLFPYQRFVEVLLLFQGGGLKALNGQGLGATSMNQALQLELRISGS 125  
DB 63 LIN-----LFVIPSGTK--QNGR-LSATTV---ATERYSLLIQISS 97

RESULT 24  
LNBOGB  
beta-galactoside-binding lectin - bovine  
N;Alternate names: 14K lectin  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: S03865; A38132; A26756  
R;Abbott, W.M.; Mellor, A.; Edwards, Y.; Feizi, T.  
Biochem. J. 259, 283-290, 1989

A;Title: Soluble bovine galactose-binding lectin: cDNA cloning reveals the complete amin  
A;Reference number: S03865; MUID:89246317; PMID:2470348  
A;Accession: S03865  
A;Molecule type: mRNA  
A;Residues: 1-135 <ABB>  
A;Cross-references: UNIPROT:P11116; EMBL:X14330; NID:9517; PIDN:CMA32508.1; PID:9518  
R;Tracey, B.M.; Feizi, T.; Abbott, W.M.; Carruthers, R.A.; Green, B.N.; Lawson, A.M.  
J. Biol. Chem. 267, 10342-10347, 1992  
A;Title: Subunit molecular mass assignment of 14,654 Da to the soluble beta-galactoside-b  
xidative inactivation.  
A;Reference number: A38132; MUID:92268070; PMID:1587821  
A;Accession: A38132  
A;Molecule type: protein  
A;Residues: 2-135 <TRA>  
A;Experimental source: heart  
A;Note: sequence extracted from NCBI backbone (NCBIP:103655)  
R;Southan, C.; Aitken, A.; Childs, R.A.; Abbott, W.M.; Feizi, T.  
FEBS Lett. 214, 301-304, 1987  
A;Title: Amino acid sequence of beta-galactoside-binding bovine heart lectin. Member of  
A;Reference number: A26756; MUID:87190976; PMID:3569527  
A;Accession: A26756  
A;Molecule type: protein  
A;Residues: 'L',12-16, 'E',18-42, 'Y',44-57, 'L',59-60, 'S',62-64, 'K',66-88, 'E',90-91, 'S',93-  
A;Experimental source: heart  
A;Note: residues 17, 43, 61, 65, 89, 92, and 131 have been revised in reference S03865  
C;Superfamily: beta-galactoside-binding lectin  
C;Keywords: acetylated amino end; lectin  
F;2-135/Product: beta-galactoside-binding lectin #status experimental <MAT>  
F;2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental  
F;3-131,17-89,43-61/Disulfide bonds: (in inactive form) #status experimental

Query Match 10.6%; Score 72; DB 1; Length 135;  
Best Local Similarity 27.3%; Pred. No. 2.7;  
Matches 35; Conservative 17; Mismatches 56; Indels 20; Gaps 6;

QY 17 LSPGVIIVRGLVLQEPKHPTVSL-RDQ-----AAHAPV-TLRASPADRTLAWI 63  
DB 12 LKPGECILVRGEVAADAKSFLNLGKDDNNLCLHFNPRFNAGDVNTVCNSKD----A 66  
QY 64 SRGQKLLISAPLPFPQRFVEVLLLFQGGGLKALNGQGLGATSMNQALQLELRIS 123  
DB 67 GAWGAEQRSEA-FPFGSGSVVEVCISFNQTDLTIKL-PDGYEPKFPNRLNLEAINVLSAG 124  
QY 124 GSVOLYCV 131  
DB 125 GDFKIKCV 132

RESULT 25  
AG0052  
probable pectinesterase (EC 3.1.1.11) [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AG0052  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
den-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AG0052  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-361 <KUR>  
A;Cross-references: UNIPROT:Q8ZIR5; GB:AL590842; PIDN:CAC89282.1; PID:gl5978520; GSPDB:G  
C;Genetics:  
A;Gene: YPO424  
C;Superfamily: pectinesterase  
C;Keywords: carboxylic ester hydrolase

Query Match 10.6%; Score 72; DB 2; Length 361;  
Best Local Similarity 27.2%; Pred. No. 8.1;  
Matches 34; Conservative 7; Mismatches 48; Indels 36; Gaps 7;

Qy 14 PQGL-----SPGVIIIRGLVLOEPKHTVSLRD-----QAAHAPVTLRASPADR 58  
Db 239 PYGLVLSRLTKPG--VPAKSFALGRPHWPTTAFDDGRVADPAALQGVAFINTVMDH 296  
Qy 59 TLAWISRWGKKLISAPFLFYPO--RPFEVLLLFQEGGLKLALNGOGLGA---TSMNQQA 113  
Db 297 IYGDNRMSGDKQ-GEKWFYFQDSRFFEA-----NNQPGGABINEGRRLS 342  
Qy 114 LEQLR 118  
Db 343 AEQLK 347

RESULT 26  
LNHUGB  
Galactin 1 [validated] - human  
A/Alternate names: 14k lectin; brain lectin; galaptin; laminin-binding protein HLBPI4; s  
C/Species: Homo sapiens (man)  
C/Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 09-Jul-2004  
A/Accession: A37134; S04330; S03866; A40299; I55292; A31038; JC1092; S04998; S24167; C26  
R/Gitt, M.A.; Barondes, S.H.  
Biochemistry 30, 82-89, 1991  
A/Title: Genomic sequence and organization of two members of a human lectin gene family.  
A/Reference number: A37134; MUID:91105104; PMID:1988031  
A/Accession: A37134  
A/Molecule type: DNA  
A/Residues: 1-135 <GIT>  
A/Cross-references: UNIPROT:P09382; GB:J05303; GB:M57678; NID:G184227; PIDN:AAB00777.1;  
A/Experimental source: Clone 182  
R/Hirabayashi, J.; Ayaki, H.; Soma, G.I.; Kasai, K.I.  
Biochim. Biophys. Acta 1008, 85-91, 1989  
A/Title: Cloning and nucleotide sequence of a full-length cDNA for human 14 kDa beta-gal  
A/Reference number: S04330; MUID:85247452; PMID:2719964  
A/Accession: S04330  
A/Molecule type: mRNA  
A/Residues: 1-135 <HRI1>  
A/Cross-references: EMBL:X14829; NID:G34342; PIDN:CAA32938.1; PID:G34343  
R/Abbott, W.M.; Feizi, T.  
Biochem. J. 259, 291-294, 1989  
A/Title: Evidence that the 14 kDa soluble beta-galactoside-binding lectin in man is enco  
A/Reference number: S03866; MUID:89246318; PMID:2719646  
A/Accession: S03866  
A/Molecule type: mRNA  
A/Residues: 1-135 <AB>  
A/Cross-references: EMBL:X15256; NID:G23238; PIDN:CAA33328.1; PID:G4377694  
R/Allen, H.J.; Gottstine, S.; Sharma, A.; DiCiccio, R.A.; Swank, R.T.; Li, H.  
Biochemistry 30, 8904-8910, 1991  
A/Title: Synthesis, isolation, and characterization of endogenous beta-galactoside-bindi  
A/Reference number: A40299; MUID:91363388; PMID:1888747  
A/Accession: A40299  
A/Molecule type: mRNA  
A/Residues: 7-66; 85-115; 119-127 <ALL>  
A/Cross-references: GB:J05338  
R/Couraud, P.  
J. Biol. Chem. 264, 1310-1316, 1989  
A/Title: Molecular cloning, characterization, and expression of a human 14-kDa lectin.  
A/Reference number: I55292; MUID:89093123; PMID:2910656  
A/Accession: I55292  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-135 <RES>  
A/Cross-references: GB:J04456; NID:G187109; PIDN:AAA36170.1; PID:G307122  
R/Hirabayashi, J.; Kasai, K.  
J. Biochem. 104, 1-4, 1988  
A/Title: Complete amino acid sequence of a beta-galactoside-binding lectin from human pl  
A/Reference number: A31038; MUID:89123203; PMID:3065332  
A/Accession: A31038  
A/Molecule type: protein  
A/Residues: 2-135 <HIR2>  
A/Experimental source: placenta  
R/Wang, Y.J.  
Prog. Biochem. Biophys. 20, 435-436, 1993

A/Title: Human brain lectin (HBL).  
A/Reference number: JC1092  
A/Accession: JC1092  
A/Molecule type: protein  
A/Residues: 2-135 <WAN>  
A/Experimental source: brain  
R/Hirabayashi, J.; Ayaki, H.; Soma, G.I.; Kasai, K.I.  
FEBS Lett. 250, 161-165, 1989  
A/Title: Production and purification of a recombinant human 14 kDa beta-galactoside-bind  
A/Reference number: S04998; MUID:89325559; PMID:2666155  
A/Accession: S04998  
A/Molecule type: protein  
A/Residues: 2, X, 4-16, X', 18-40 <HIR>  
A/Experimental source: recombinant protein expressed in E. coli  
R/Castronovo, V.; Luyten, F.; van den Brule, F.; Sobel, M.E.  
Arch. Biochem. Biophys. 297, 132-138, 1992  
A/Title: Identification of a 14-kDa laminin binding protein (HLBP14) in human melanoma c-  
A/Reference number: S24167; MUID:92344405; PMID:1386213  
A/Accession: S24167  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 20-29; 50-64; 65-74; 132-135 <CAS>  
R/Gitt, M.A.; Barondes, S.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 7603-7607, 1986  
A/Title: Evidence that a human soluble beta-galactoside-binding lectin is encoded by a f  
A/Reference number: A94130; MUID:87016997; PMID:3020551  
A/Accession: C26495  
A/Molecule type: protein  
A/Residues: 22-42, X', 44-66, 70, XX', 73, 75-78, Y', 80, X', 82; 101-102, X', 104-105, Q', 107-1  
R/Hirabayashi, J.; Kawasaki, H.; Suzuki, K.; Kasai, K.  
J. Biochem. 101, 987-995, 1987  
A/Title: Further characterization and structural studies on human placenta lectin.  
A/Reference number: A26761; MUID:87279994; PMID:3611046  
A/Accession: A26761  
A/Molecule type: protein  
A/Residues: 70-84, X', 86-87; 122-135 <HI2>  
A/Experimental source: placenta  
C/Comment: The lectin is inactivated upon the formation of disulfide bonds.  
C/Genetics:  
A/Gene: LGLALS1  
A/Cross-references: GDB:I26889; OMIM:150570  
A/Map position: 22q12-22q13.1  
C/Superfamily: beta-galactoside-binding lectin  
C/Keywords: acetylated amino end; glycoprotein; lectin  
F:2-135/Product: beta-galactoside-binding lectin #status experimental <MAT>  
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted  
F:96/Binding site: carbohydrate (Asn) (covalent) #status absent  
Query Match 10.4%; Score 71; DB 1; Length 135;  
Best Local Similarity 25.2%; Pred. No. 3.4;  
Matches 31; Conservative 18; Mismatches 64; Indels 10; Gaps 4;  
Qy 17 LSPGVIIIRGLVLOEPKHTVSL-----RQAAHAPVTLRASPADRTLAWISR----WGQ 68  
Db 12 LKPGECILRVGEVADPAKSFVILNGKDSNNLCLHFNFRNAGDANIVCNCKGGAWGT 71  
Qy 69 KKLISAPFLFYPOFFFEVLLLFQEGGLKLALNGOGLGATSMNQQALEQLRELIRSGSVOL 128  
Db 72 EQR-EAVFPFGSGVAEVCITFDQANLTVKL-PDGYEFKFPNRLNLEAINYMAADGDFKI 129  
Qy 129 YCV 131  
Db 130 KCV 132  
RESULT 27  
C75570  
chromosome partitioning protein - Deinococcus radiodurans (strain R1)  
C/Species: Deinococcus radiodurans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
A/Accession: C75570  
R/White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M

```

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75570
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-288 <WHI>
A;Cross-references: UNIPROT:Q9RYD8; GB:AE001865; GB:AE000513; NID:g6457669; PIDN:AAF0560
C;Genetics:
A;Gene: DR0012
A;Map position: 1
C;Superfamily: Bacillus subtilis transport protein spo0J

Query Match      10.4%; Score 71; DB 2; Length 288;
Best Local Similarity 24.6%; Pred. No. 8;
Matches 33; Conservative 18; Mismatches 59; Indels 24; Gaps 4;

Qy      15  QGLSPGQVIIVRGVLQEPKHPKTVSLRDQ---AAHAPVTLRASFADRTLAW----- 62
Db      150 QAVKGKGRSTVTNALRLTLPEPVLRLDEGISASHARAVLTQPEADRLWAFQIRSRGL 209
Qy      63  -----ISRWGQKKLIAPFLFPQRPFEVL--LFOEGGLKALANGQGLGATSMNQ 111
Db      210 NVREAEAALKERGRDQKGQAPIKVNPFPRAVRQLELDLSERTGTRVKITGEDKGRVELNY 269
Qy      112 QALEQL--RELRI8 124
Db      270 GSRELDRILOILG 283

```

RESULT 28

F72632

probable myo-inositol-1-phosphate synthase APE1517 - Aeropyrum pernix (strain KI)

C/Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 16-Aug-2004

C/Accession: F72632

R/Kawakabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kanao, R. 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain KI

A/Reference number: A72450; MUID:99310339; PMID:10382966

A/Accession: F72632

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-392 <KAW>

A/Cross-references: UNIPROT:Q9VBT2; DBJ:AP000061; NID:G5104821; PIDN:EAA80516.1; PID:dl100000000

A/Experimental source: strain KI

C/Genetics:

A/Gene: APE1517

C/Superfamily: Myo-inositol-1-phosphate synthase

```

Query Match      10.4%; Score 71; DB 2; Length 392;
Best Local Similarity 27.5%; Pred. No. 11;
Matches 30; Conservative 14; Mismatches 57; Indels 8; Gaps 2;

Qy 15 QGLSPGOVIVRGLVLOEPTVSLDDQAHAHPVTLRASFADRTLAWISRWQK----K 70
Db 124 QHLRPDVIVNITTEAAPEPTLESVRDRRGIVSASQAYALAVLRVYAEAREARVALVN 183

Qy 71 LISAPFLFYQRFVEFVLLLFQEGGLKIALNGQGLGATSMNQQALEQIRE 119
Db 184 AIPAPLANDP-----VLVSMFEERAGSLLLGGDGGATGATPTADLLEHIAE 228

```

RESULT 29  
E83188  
1-deoxy-d-xylose 5-phosphate reductoisomerase PA3650 [imported] - *Pseudomonas aeruginosa*  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: E83188  
R:Stover, C. K.; Pham, X. O.; Erwin, A. L.; Mizoguchi, S. D.; Warren, P.; Hickey, M. J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406. 959-964, 2000  
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: E83188  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-396 <STO>  
A;Cross-references: UNIPROT:Q9RGU6; GB:AE004785; GB:AE0049809; PIDN:AAG0703  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: dxr; PA3650  
C;Superfamily: 1-deoxy-D-xylulose 5-phosphate reductoisomerase

	Query Match	10.4%;	Score 71;	DB 2;	Length 396;	
	Best Local Similarity	24.1%;	Pred. No. 12;			
	Matches	32;	Conservative 19;	Mismatches 50;	Indels 32;	Gaps 5;
Qy	2 SPRELVPCSHA-----LPGLSPGGVIIVRGLVLOEP---KHPTVSLRDQAAH----AP 48	:	:	:	:	:
Db	271 NPMRTPISYAMAWPERIDSGVSPLOWFAVGRLDFFORPDEQRFPCLRASQAQAE TGSGAP 330	:	:	:	:	:
Qy	49 VTLRAGFADRTLAWISRWGOKKLISAPPFYPPFRFEVLILLFOGGKILNALNCGQLGATS 108	:	:	:	:	:
Db	331 AMLNAA-----NEVAVAFLERHIFSDIAVIED-----VLNEEAVTAVE 371	:	:	:	:	:
Qy	109 MNQQALEURELELR 121	:	:	:	:	:
Db	372 SLDOVLAADRARR 384	:	:	:	:	:

RESULT 30  
AI2811  
serine proteinase htrA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C/Accession: AI2811  
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.  
ster, E.W.  
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A/Reference number: AB2577; MUID:21608550; PMID:11743193  
A/Accession: AI2811  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-468 <KUR>  
A/Cross-references: UNIPROT:Q8UE46; GB:AE008688; PIDN:AAL42911.1; PID:gl7740366; GSPDB:GN  
A/Experimental source: strain C58 (Dupont)  
C/Genetics:  
A/Gene: htrA  
A/Map position: circular chromosome  
C/Superfamily: Helicobacter serine proteinase

	Query Match	10.4%	Score 71;	DB 2;	Length 468;
	Best Local Similarity	25.0%;	Pred. No. 14;		
	Matches	36;	Conservative	20;	Mismatches
				38;	Indels
				50;	Gaps
					6;
Qy	16 GLSPGQVI-IVRGVLQEP-----KXPTVSLRDQAHAAP	48			
	: ::   :	:	:	:	:
Dd	309 GIEPQQVTAVNGLEVEHPDALGYRLTTAGIGSAELTVVEKEGEKKLTALTATETAP	368			
	: ::   :	:	:	:	:
Qy	49 -----VTLRASFDRTLIAMISRWGQKKLISAPFLFYPPQRFEVLL-----FQ	91			
	: ::   :	:	:	:	:
Dd	369 RDERELLEGRNPFAGATVANLS----PKL--ADELRMPQSQTGVVITDVKEGSPAYRVGFQ	422			
	: ::   :	:	:	:	:
Qy	92 EGGKLIALNGOGLGATSMNQALE	115			
	: ::   :	:	:	:	:
Dd	423 PKDVILSLNGADICGSTAAVEKALD	446			
	: ::   :	:	:	:	:



Search completed: October 17, 2005, 08:42:51  
Job time : 20 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:29:48 ; Search time 179 Seconds  
(without alignments)  
380.483 Million cell updates/sec

Title: HJACE54  
Perfect score: 682  
Sequence: 1 msprlevpcshalpqlspg.....leglreirsgsvqlcvs 133

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:  
1: uniprot\_eprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	671	98.4	296	2 Q9NZ03	Q9NZ03 homo sapien
2	671	98.4	336	1 LEGC HUMAN	Q96dt0 homo sapien
3	556	81.5	314	1 LEGC_MOUSE	Q91vd1 mus musculus
4	482	70.7	383	1 Q8C3K0	Q8C3K0 mus musculus
5	146.5	21.5	171	2 Q6NVI7	Q6NVI7 xenopus tro
6	143.5	21.0	171	2 Q6DDR8	Q6DDR8 xenopus lae
7	142.5	20.9	171	2 Q68FJ4	Q68FJ4 xenopus lae
8	134	19.6	315	2 Q7ZTB5	Q7ZTB5 xenopus lae
9	120.5	17.7	172	2 Q6P5T6	Q6P5T6 homo sapien
10	120.5	17.7	172	2 Q8VED9	Q8VED9 mus musculus
11	119	17.4	144	1 LEGS_RAT	P47967 rattus norv
12	118.5	17.4	172	2 Q9P005	Q9P005 homo sapien
13	109.5	16.1	316	1 LEGH HUMAN	Q00214 homo sapien
14	108.5	15.9	261	2 Q7ZSY1	Q7ZSY1 xenopus lae
15	102	15.0	322	2 Q9SL83	Q9SL83 mus musculus
16	102	15.0	353	1 LEG9_MOUSE	Q08573 mus musculus
17	99	14.5	323	2 Q6QZP2	Q6QZP2 canis fami
18	98	14.4	322	2 Q6P7Q6	Q6P7Q6 rattus norv
19	98	14.4	326	1 LEG4_MOUSE	Q8K419 mus musculus
20	98	14.4	354	1 LEG9_RAT	P97840 rattus norv
21	97.5	14.3	165	2 Q7TPX9	Q7TPX9 mus musculus
22	97	14.2	308	2 Q7ZTB8	Q7ZTB8 xenopus lae
23	96	14.1	301	1 LEG6_MOUSE	O54891 mus musculus
24	96	14.1	332	2 Q8QGD9	Q8QGD9 gallus gall
25	95.5	14.0	135	1 LEG7_HUMAN	P47929 homo sapien
26	94.5	13.9	727	2 Q44230	Q44230 anabaena sp
27	94.5	13.9	727	2 Q7A2H2	Q7A2H2 anabaena sp
28	94	13.8	204	2 Q7SYPO	Q7SYPO xenopus lae
29	94	13.8	319	2 Q7ZTB6	Q7ZTB6 xenopus lae
30	93.5	13.7	218	2 Q816N9	Q816N9 anopheles s
31	93.5	13.7	316	2 Q6IN24	Q6IN24 rattus norv

32	93	13.6	262	2 Q90713	Q90713 gallus gall
33	91.5	13.4	223	2 Q8C6H0	Q8C6H0 mus musculu
34	91.5	13.4	316	1 LEG8_MOUSE	Q9J115 mus musculu
35	91.5	13.4	359	2 Q9BXC8	Q9BXC8 homo sapien
36	91	13.3	228	2 Q6TGN4	Q6TGN4 brachydanio
37	89.5	13.1	256	2 Q6P8B0	Q6P8B0 xenopus tro
38	88.5	13.0	211	2 Q7QHP5	Q7QHP5 anopheles g
39	88	12.9	334	2 Q6DDQ2	Q6DDQ2 xenopus lae
40	88	12.9	343	2 Q8UW98	Q8UW98 xenopus lae
41	87.5	12.8	300	2 Q96Q57	Q96Q57 homo sapien
42	87.5	12.8	316	1 LEG8_RAT	Q62665 rattus norv
43	87.5	12.8	359	2 Q8TEV1	Q8TEV1 homo sapien
44	87	12.8	356	2 Q6DKI2	Q6DKI2 homo sapien
45	85.5	12.5	263	1 LEG3_MOUSE	P61110 mus musculu
46	85.5	12.5	264	2 Q8C353	Q8C353 mus musculu
47	84	12.3	241	2 Q8NNB5	Q8NNB5 corynebacte
48	84	12.3	324	1 LEG4_RAT	P38552 rattus norv
49	82.5	12.1	135	1 LEG7_MOUSE	Q54974 mus musculu
50	82.5	12.1	135	1 LEG7_RAT	P97590 rattus norv
51	82.5	12.1	762	2 Q88MS6	Q88MS6 pseudomonas
52	81.5	12.0	323	1 LEG4_HUMAN	P56470 homo sapien
53	81.5	12.0	340	2 Q8UW99	Q8UW99 xenopus lae
54	81	11.9	847	1 STA6_HUMAN	P42226 homo sapien
55	81	11.9	847	2 Q71UP4	Q71UP4 homo sapien
56	80.5	11.8	136	2 Q9CRB1	Q9CRB1 m mus muscu
57	80.5	11.8	241	1 LEG3_RABIT	P47845 oryctolagus
58	80.5	11.8	558	2 Q7VX36	Q7VX36 bordetella
59	80.5	11.8	558	2 Q7W722	Q7W722 bordetella
60	80.5	11.8	558	2 Q7WLD3	Q7WLD3 bordetella
61	80	11.7	317	2 Q9XSM8	Q9XSM8 sus scrofa
62	80	11.7	349	2 Q9XSM9	Q9XSM9 sus scrofa
63	79	11.6	194	2 Q9BRJ6	Q9BRJ6 homo sapien
64	78.5	11.5	136	2 Q99ML7	Q99ML7 mus musculu
65	78.5	11.5	197	1 Y105_HABDU	Y105KR7 haemophilus
66	78.5	11.5	261	1 LEG3_RAT	P08699 rattus norv
67	78.5	11.5	323	1 LEG4_PIG	Q29058 sus scrofa
68	78.5	11.5	337	2 Q83WT5	Q83WT5 streptomyce
69	78	11.4	310	2 Q7T118	Q7T118 brachydanio
70	78	11.4	320	2 Q6DHG3	Q6DHG3 brachydanio
71	78	11.4	320	2 Q6TLH6	Q6TLH6 brachydanio
72	78	11.4	320	2 Q7T119	Q7T119 brachydanio
73	77.5	11.4	136	2 Q9D7K4	Q9D7K4 mus musculu
74	77.5	11.4	244	1 LEG3_CRILLO	P47953 cricetus
75	77.5	11.4	249	1 LEG3_HUMAN	P17931 homo sapien
76	77.5	11.4	250	2 Q6FGL0	Q6FGL0 homo sapien
77	77.5	11.4	250	2 Q61BA7	Q61BA7 homo sapien
78	77.5	11.4	328	2 Q9TUB8	Q9TUB8 oryctolagus
79	77.5	11.4	727	2 Q7S218	Q7S218 neurospora
80	77	11.3	320	2 Q8TGN3	Q8TGN3 brachydanio
81	77	11.3	348	2 Q8QGB1	Q8QGB1 oncorhynch
82	77	11.3	530	2 Q8C1W2	Q8C1W2 mus musculu
83	77	11.3	1293	2 Q8MUV2	Q8MUV2 giardia lam
84	77	11.3	1293	2 Q7QSN9	Q7QSN9 giardia lam
85	76.5	11.2	250	2 Q6NVH9	Q6NVH9 homo sapien
86	76.5	11.2	295	1 LEG3_CANFA	P38486 canis fami
87	76	11.1	311	2 Q8WYQ7	Q8WYQ7 homo sapien
88	76	11.1	355	1 LEG9_HUMAN	O00182 homo sapien
89	75	11.0	136	2 Q6J275	Q6J275 ovis aries
90	74.5	10.9	126	2 Q7R579	Q7R579 giardia lam
91	74.5	10.9	277	2 Q9NJV1	Q9NJV1 haemonchus
92	74.5	10.9	283	2 Q86FX2	Q86FX2 haemonchus
93	74.5	10.9	283	2 Q9NUV0	Q9NUV0 haemonchus
94	74.5	10.9	373	2 Q9DCS3	Q9DCS3 mus musculu
95	74.5	10.9	373	2 Q99L39	Q99L39 mus musculu
96	74.5	10.9	555	2 Q44WR0	Q44WR0 bacteroides
97	74.5	10.9	703	2 Q8DASS	Q8DASS vibrio vuln
98	74	10.9	209	2 Q9KHV9	Q9KHV9 rhizobium m
99	74	10.9	652	2 Q6N349	Q6N349 rhodospesdo
100	73.5	10.8	283	1 LEG1_HAECO	O44126 haemonchus

ALIGNMENTS

```
RESULT 1
Q9NZ03 Q9NZ03 PRELIMINARY; PRT; 296 AA.
AC Q9NZ03;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin-related inhibitor of proliferation isoform b.
GN Name=GRIP1;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Liver;
RA Yang R.-Y., Hsu D.K., Ni J., Yu L., Liu F.-T.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AF222694; AAF34676.1; -.
DR HSP; P17931; IAK3.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLEGT; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW Galectin; Lectin.
SQ SEQUENCE 296 AA; 33257 MW; 1A160DA2E9A48ACC CRC64;

Query Match 98.4%; Score 671; DB 2; Length 296;
Best Local Similarity 99.2%; Pred. No. 2.4e-64;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLVPCSHALPQGLSPQGVIIIRGLVLQEPKFTVSRDQAAHAPVTLRASFADRTL 60
DB 164 MSRLVPCSHALPQGLSPQGVIIIRGLVLQEPKFTVSRDQAAHAPVTLRASFADRTL 223
QY 61 AWISRWQKKLISAPPLFYQRFVFLVLLPQEGGLKLALNGQGLGATSMNQALEQRL 120
DB 224 AWISRWQKKLISAPPLFYQRFVFLVLLPQEGGLKLALNGQGLGATSMNQALEQRL 283
QY 121 RISGSVOLYCVHS 133
DB 284 RISGSVOLYCVHS 296

RESULT 2
LEGC_HUMAN STANDARD; PRT; 336 AA.
AC Q96DT0; Q96DS9; Q96PR9; Q9H258; Q9NZ02;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galectin-12 (Galectin-related inhibitor of proliferation).
GN Name=LGALS12; Synonyms=GRIP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D), AND CHARACTERIZATION.
RC TISSUE=Adipose tissue;
RA MEDLINE=21433969; PubMed=11435439; DOI=10.1074/jbc.M105097200;
RA Hotta K., Funahashi T., Matsukawa Y., Takahashi M., Nishizawa H.,
RA Kishida K., Matsuda M., Kuriyama H., Kihara S., Nakamura T.,
RA Tochino Y., Bodkin N.L., Hansen B.C., Matsuzawa Y.;
RT "Galectin-12, an adipose-expressed galectin-like molecule possessing
RT apoptosis-inducing activity."
RL J. Biol. Chem. 276:34089-34097(2001).
RN [2]

SEQUENCE FROM N.A. (ISOFORMS B; E AND F).
TISSUE=Retina;
MEDLINE=21283005; PubMed=11283015; DOI=10.1074/jbc.M010914200;
Yang R.-Y., Hsu D.K., Yu L., Ni J., Liu F.-T.;
"Cell cycle regulation by galectin-12, a new member of the galectin
superfamily."
J. Biol. Chem. 276:20252-20260(2001).
[3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Blood;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL -1- FUNCTION: Binds lactose. May participate in the apoptosis of
CC adipocytes.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=A;
CC IsoId=Q96DT0-1; Sequence=Displayed;
CC Name=B; Synonyms=GRIP1a; Sequence=VSP_003100;
CC IsoId=Q96DT0-2; Sequence=VSP_003100;
CC Name=C;
CC IsoId=Q96DT0-3; Sequence=VSP_003102;
CC Name=D;
CC IsoId=Q96DT0-4; Sequence=VSP_003100, VSP_003102;
CC Name=E; Synonyms=1;
CC IsoId=Q96DT0-5; Sequence=VSP_003099;
CC Name=F; Synonyms=2;
CC IsoId=Q96DT0-6; Sequence=VSP_003099, VSP_003101;
CC -1- TISSUE SPECIFICITY: Not widely expressed. Predominantly expressed
CC in adipose tissue.
CC -1- DOMAIN: Contains two homologous but distinct carbohydrate-binding
CC domains.
CC -1- SIMILARITY: Belongs to the galectin (galectin/S-lectin) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF244974; AAK77328.1; -.
CC EMBL; AF244975; AAK77329.1; -.
CC EMBL; AF244976; AAK77330.1; -.
CC EMBL; AF244977; AAK77331.1; -.
CC EMBL; AF310686; AAG40863.1; -.
CC EMBL; AF310687; AAG40864.1; -.
CC EMBL; AF222695; AAF34677.1; -.
CC EMBL; BC028222; AAH28222.1; -.
CC HSP; P17931; IAK3.
CC Genew; HGNC:15788; LGALS12.
CC H-InvDB; HIX0009737; -.

```

```

DR MIM; 606096; --
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLCT; 1
DR PROSITE; PS00309; GALACTIN; FALSE_NEG.
KW Alternative splicing; Apoptosis; Galectin; Lectin; Nuclear protein;
KW Repeat.
FT DOMAIN 47 183 Galaptin 1.
FT DOMAIN 194 336 Galaptin 2.
FT VARSPLIC 1 22 Missing (in isoform E and isoform F).
FT VARSPLIC 1 61 Missing (in isoform B and isoform D).
FT VARSPLIC 74 74 Missing (in isoform F).
FT VARSPLIC 200 208 Missing (in isoform C and isoform D).
FT VARSPLIC 336 AA; 37541 MW; F35D0A0CE503E795 CRC64;
SQ SEQUENCE 98.4%; Score 671; DB 1; Length 336;
Query Match 99.2%; Pred. No. 2.8e-64;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MSPLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSLRDQAHAHPVTLRASFAFRTL 60
Db 204 MSPLEVPCHALPQGLSPGVIIIRGLVLPQPKHPTVSLRDQAHAHPVTLRASFAFRTL 263
Qy 61 AWISRWGQKLLISAPFLYFPPRFEVLLIFQEGGLKALNGOGLGATSMNQALEQREL 120
Db 264 AWISRWGQKLLISAPFLYFPPRFEVLLIFQEGGLKALNGOGLGATSMNQALEQREL 323
Qy 121 RISGSVOLYCVHS 133
Db 324 RISGSVOLYCVHS 336

RESULT 3
LEGC MOUSE STANDARD; PRT; 314 AA.
AC Q91VD1; Q8CC44; Q8K2L7; Q9UKX2;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galectin-12.
GN Name=Lgal12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC STRAIN=C57BL/KsK; TISSUE=adipose tissue;
RX MEDLINE=21423969; PubMed=11435439; DOI=10.1074/jbc.M105097200;
RA Hotta K., Funahashi T., Matsukawa Y., Takahashi M., Nishizawa H.,
RA Kishida K., Matsuda M., Kuriyama H., Kihara S., Nakamura T.,
RA Tochino Y., Bodkin N.L., Hansen B.C., Matsuzawa Y.;
RT "Galectin-12, an adipose-expressed galectin-like molecule possessing
RT apoptosis-inducing activity.";
RL J. Biol. Chem. 276:34089-34097 (2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RX MEDLINE=21283005; PubMed=11283015; DOI=10.1074/jbc.M010914200;
RT Yang R.-Y., Hsu D.K., Yu L., Ni J., Liu F.-T.;
RT "Cell cycle regulation by galectin-12, a new member of the galectin
RT superfamily.";
RL J. Biol. Chem. 276:20252-20260 (2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

```

```

RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gattinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Grimmer S., Gustincich S., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kanai A., Kawai H., Kawasawa Y., Lee Y., Lenhard B., Lyons P.A.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Naglath D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nakata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takemura Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Varardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP FUNCTION: Binds lactose. May participate in the apoptosis of
RP adipocytes.
RP SUBCELLULAR LOCATION: Nuclear.
RP ALTERNATIVE PRODUCTS:
RP Event=Alternative splicing; Named isoforms=2;
RP Name=1;
RP IsoId=Q91VD1-1; Sequence=Displayed;
RP Name=2;
RP IsoId=Q91VD1-2; Sequence=VSP_010321;
RN [6]
RP DOMAIN: Contains two homologous but distinct carbohydrate-binding
RP domains.
RN [7]
RP SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
RN [8]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch)
RN [9]
RP EMBL; AF244979; AAK77327.1; -

```

```

DR EMBL: AF244978; AAK77326.1; -.
DR EMBL: AF223223; AAF34682.1; -.
DR EMBL: AK033535; BAC28345.1; -.
DR EMBL: BC030890; AAH30890.1; -.
DR HSSP: P17931; 1A3K.
DR MGD; MGI:1929094; Lgale12.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Alternative splicing; Apoptosis; Galectin; Lectin; Nuclear protein;
FT Repeat.
FT DOMAIN 25 161 Galaptin 1.
FT DOMAIN 172 314 Galaptin 2.
FT VARSPLIC 167 179 Missing (in isoform 2).
FT CONFLICT 228 228 V -> A (in Ref. 1).
FT CONFLICT 273 273 E -> K (in Ref. 4).
SQ SEQUENCE 314 AA; 35460 MW; 9C9D33A7583DD09D CRC64;

Query Match 81.5%; Score 556; DB 1; Length 314;
Best Local Similarity 81.7%; Pred. No. 8e-52;
Matches 107; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 2 SPRLVPCSHALPQGLSPGVIVRGLVLOEPKHPTVSLRDQAHAAPVTLRASFAFDRTLA 61
DB 183 SPRLVPCSHALPQGLSPGVIVRGLVLOEPKHPTVSLRDQAHAAPVTLRASFAFDRTLA 242
QY 62 WISRWGOKLISAPFLFYPPQRFVFFVLLLFQGGKLGALNQGGLGATSMNQALEQ 121
DB 243 WVSWSGRKKLISAPFLFHPQRFVFFVLLLFQGGKLGALNQGGLGATSLDQKALEQ 302
QY 122 ISGSVQLYCVH 132
DB 303 ISGNVHLVYCVH 313

RESULT 4
ID Q8C3K0 PRELIMINARY; PRT; 383 AA.
AC Q8C3K0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 10 days lactation, adult female mammary gland cDNA, RIKEN
DE full-length enriched library, clone:D73000J21 product:lectin,
DE galactose binding, soluble 12, full insert sequence.
GN Name=Lgale12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";

```

---

```

RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases
CC -!- SIMILARITY: Belongs to the galectin (galaptin/S-lectin) family.
DR EMBL; AK085660; BAC39498.1; -.
DR HSSP; P17931; 1A3K.
DR MGD; MGI:1929094; Lgals12.
DR GO; GO:0005739; C:mitochondrion; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
DR Galectin; Lectin.
KW SEQUENCE 383 AA; 43036 MW; B7C10FA5746977DF CRC64;

Query Match 70.7%; Score 482; DB 2; Length 383;
Best Local Similarity 80.9%; Pred. No. 1.1e-43;
Matches 93; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 2 SPRLVPCSHALPQGLSPGVIVRGLVLOEPKHPTVSLRDQAHAAPVTLRASFAFDRTLA 61
DB 183 SPRLVPCSHALPQGLSPGVIVRGLVLOEPKHPTVSLRDQAHAAPVTLRASFAFDRTLA 242
QY 62 WISRWGOKLISAPFLFYPPQRFVFFVLLLFQGGKLGALNQGGLGATSMNQALEQ 116
DB 243 WVSWSGRKKLISAPFLFHPQRFVFFVLLLFQGGKLGALNQGGLGATSLDQKALEQ 297

RESULT 5
ID Q6NV17 PRELIMINARY; PRT; 171 AA.
AC Q6NV17;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC76272.

```

GN Name=MGC76272;  
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Gerhard D.S.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.  
DR GO:0005529; F:sugar binding; IEA.  
DR InterPro: IPR008985; ConA like lec\_gl.  
DR InterPro: IPR001079; Galectin.  
DR Pfam: PF00337; Gal-bind lectin; 1.  
DR SMART: SM00276; GLEGT; 1.  
DR Galectin; Hypothetical protein; Lectin.  
SQ SEQUENCE 171 AA; 19111 MW; 39CF31F7C65DA040 CRC64;  
Query Match 21.5%; Score 146.5; DB 2; Length 171;  
Best Local Similarity 31.6%; Pred. No. 1.2e-07;  
Matches 43; Conservative 27; Mismatches 55; Indels 11; Gaps 5;  
QY 3 PRLEVP-CSHALPQGLSPGVIIYRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFA 56  
Db 33 PRLTVPPFCGH-IKGMRPGKILINGVNLPEKSPDIRLTCGSDSDPAADIAIELRAEFA 91  
QY 57 DRTL---AWIS-RWGQKLLISAPFLFYQRFQRFVLLLFQEGGLKALNGOGLGATSMNQ 112  
Db 92 DKQLLRNACVSGKMGEESSAIPYFPFIADQPRVRIHCEHPRLRIFVDGHQFLDFYHRVE 151  
QY 113 ALEQLRELRLISGSVOL 128  
Db 152 TSLAINTIKINGDLQL 167  
RESULT 6  
Q6DDR8 PRELIMINARY; PRT; 171 AA.  
AC Q6DDR8;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE MGC82421 protein.  
GN Name=MGC82421;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
initiative";  
RL Dev. Dyn. 225:384-391 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Gerhard D.S.;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.  
DR EMBL: BC077459; AAH77459.1; --  
DR GO:0005529; F:sugar binding; IEA.  
DR InterPro: IPR008985; ConA like lec\_gl.  
DR InterPro: IPR001079; Galectin.  
DR Pfam: PF00337; Gal-bind lectin; 1.  
DR SMART: SM00276; GLEGT; 1.  
DR Galectin; Lectin.  
SQ SEQUENCE 171 AA; 19233 MW; 7FE19E3C7471D402 CRC64;  
Query Match 21.0%; Score 143.5; DB 2; Length 171;  
Best Local Similarity 30.9%; Pred. No. 2.5e-07;  
Matches 42; Conservative 29; Mismatches 54; Indels 11; Gaps 5;  
QY 3 PRLEVP-CSHALPQGLSPGVIIYRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFA 56  
Db 33 PRLTVPPFCGH-IKGMRPGKILINGVNLPEKSPDIRLTCGSDSDPAADIAIELRAEFS 91  
QY 57 DRTL---AWIS-RWGQKLLISAPFLFYQRFQRFVLLLFQEGGLKALNGOGLGATSMNQ 112  
Db 92 DKQFLRNARVSGKMGEESSAIPYFPFIADQPRVRIHCEHPRLRIFVDGHQFLDFYHRVE 151  
QY 113 ALEQLRELRLISGSVOL 128  
Db 152 TSLAINTIKINGDLQL 167  
RESULT 7  
Q68FJ4 PRELIMINARY; PRT; 171 AA.  
ID Q68FJ4  
AC Q68FJ4;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)





```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pooled;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; BC062691; AA62691.1; -.
DR HSP; P17931; 1A3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Galectin; Hypothetical protein; Lectin.
SQ SEQUENCE 172 AA; 18986 MW; E2D41EA68EF44DEC CRC64;
Query Match 17.7%; Score 120.5; DB 2; Length 172;
Best Local Similarity 29.4%; Pred. NO. 8.1e-05;
Matches 40; Conservative 27; Mismatches 58; Indels 11; Gaps 5;

Qy 3 PRLVPP-CSHALPQGLSPGVIIVRGLVLPQPKHPTVSL-----RDQAAHAPVTLRASFA 56
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 PRLVPPCGH-IKGMRPKKVLWGVLDLNPSPFSLTCDGSDPPADVAIELKAVFT 92
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 57 DRTL---AWIS-RWGQKXLSAPFLYFPQRFVLLPQEGGLKALNGQGLGATSMNQ 112
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 DQLLRNSCISGERGEQSAIPYFPFIPDQPFVRIILCEHPRFVFDGHQLDFPYHRIQ 152
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 113 ALEQLRELRLISGSVQL 128
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 TLSAIDTIKINGDLQI 168
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q8VED9 Q8VED9 PRELIMINARY; PRT; 172 AA.
ID Q8VED9
AC Q8VED9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE RIKEN CDNA 1110087D22 (Mus musculus adult male aorta and vein cDNA,
DE RIKEN full-length enriched library, clone:A530071M23
DE product:hypothetical Galectins (previously S-lectins) containing
DE protein, full insert sequence).
GN Name=1110067D22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

```

```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto K., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Adachi J., Azawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanganaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

```

```

RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJJ databases.
CC -!- SIMILARITY: Belongs to the galectin (galaptin/S-lectin) family.
DR ENBL; BC019131; AAH19131.1; -.
DR DR ENBL; AK041053; BAC30800.1; -.
DR HSSP; P17931; IA3K.
DR GMD; MG1.1916114; L110067D22Rik.
DR GO; GO:0005529; F:sugar binding; IEA.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SW00276; GLECT; 1.
KW Galectin; Hypothetical protein; Lectin.
SQ SEQUENCE 172 AA; 18956 MW; 56D40BE7DS6FC336 CRC64;

Query Match      17.7%; Score 120.5; DB 2; Length 172;
Best Local Similarity 29.4%; Pred.No. 8.le-05;
Matches 40; Conservative 27; Mismatches 58; Indels 11; Gaps 5;

Qy    3 PRLEVP-CSHALPGQLSPGVQIIIVRGVLVLPQPKHTVSL-----RDQAHAHPVTIRASFA 56
       ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    34 PRLVFPFGCH-IKGMRPGKKVLWVGIVDLNPFESFALSITCGDSEDPADVAIELKAFT 92

Qy    57 DRTL---AWIS-RWGQKLISAPFLPPQRFFEVLLVFQEGGLKIALNGQIGATSMNQ 112
       ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    93 DRQLLRNCISCIGERGESSQAIPFPFDQPFRVEILCEHPRFRVFDGHOLFDFYHRIQ 152

Qy    113 ALGQLRELRISSVOL 128
       ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    153 TLSAIDTIKINGDLQI 168

RESULT 11
LEG5_RAT          STANDARD;          PRT;   144 AA.
ID LEG5_RAT
AC P47967;
DT 01-FEB-1996 (Rel. 33, Created)
DI 01-FEB-1996 (Rel. 33, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Galectin-5 (RL-18).
GN Name=lgals5;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSU=Reticulocytes;
RX MEDLINE=95197487; PubMed=7890611; DOI=10.1074/jbc.270.10.5032;
RA Gitt M.A., Wiser M.P., Leffler H., Herrmann J., Xia Y.-R., Massa S.M.,
RA Cooper D.N.W., Lusis A.J., Barondes S.H.;
RT "Sequence and mapping of galectin-5, a beta-galactoside-binding
RT lectin, found in rat erythrocytes.";
RL J. Biol. Chem. 270:5032-5038(1995).
CC -!- FUNCTION: May function in erythrocyte differentiation.
CC -!- SUBUNIT: Monomer.
CC -!- TISSUE SPECIFICITY: Erythrocytes.
CC -!- SIMILARITY: Belongs to the galectin (galaptin/S-lectin) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; L36862; AAC42050.1; -.
DR ENBL; I27111; AAA65445.1; -.
DR PIR; A55932; A55932.
DR HSSP; P47929; LBKZ.
DR RGD; 3004; Lgale5.
DR InterPro; IPRO08985; ConA like lec ql.
```

Oy 113 ALEQRLBELRISGSVOL 128  
 Db 153 TLSAIDRIKINGDLQI 168

RESULT 13  
 LEG8 HUMAN STANDARD; PRT; 316 AA.  
 AC \_O02I4; O15215; Q96B92; Q9H584; Q9H585; Q9UEZ6; Q9UP32; Q9UP33;  
 AC Q9UP34;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-JAN-2005 (Rel. 46, Last annotation update)  
 DE Galectin-8 (Gal-8) (prostate carcinoma tumor antigen 1) (PCTA-1) (Po66  
 DE carbohydrate-binding protein) (Po66-CBP).  
 GN Name=LGAL8;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Prostate;  
 RC MEDLINE=96293510; PubMed=8692978; DOI=10.1073/pnas.93.14.7252;  
 RA Su Z.-Z., Lin J., Shen R., Fisher P.B., Goldstein N.I., Fisher P.B.;  
 RT "Surface-epitope masking and expression cloning identifies the human  
 RT prostate carcinoma tumor antigen gene PCTA-1 a member of the galectin  
 RT gene family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:7252-7257 (1996).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Hippocampus;  
 RC Hadari Y.R., Eisenstein M., Zakut R., Zick Y.;  
 RT "Galectin-8: on the road from structure to function.";  
 RL Trends Glycosci. Glycotechnol. 9:103-112 (1997).  
 [3]  
 RP SEQUENCE FROM N.A.; AND ALTERNATIVE SPLICING.  
 RP TISSUE=Lung carcinoma;  
 RC Brichory F., Biron N., Desrues B., Bourguet P., Le Pennec J.P.,  
 RA Dazord L.;  
 RT "Molecular cloning of a beta-galactoside-binding lectin related to  
 RT galectin-8 and identified in human lung carcinoma.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RP Maier C., Haussler J., Roesch K., Moschagath E., Vogel W.;  
 RT "Genomic organization and expression of the human galectin-8 gene.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20438187; PubMed=10980616; DOI=10.1038/sj.onc.1203767;  
 RA Gopalakrishnan R.V., Roberts T., Tuli S., Kang D., Christiansen K.A.,  
 RA Fisher P.B.;  
 RT "Molecular characterization of prostate carcinoma tumor antigen-1,  
 RT pCTA-1, a human galectin-8 related gene.";  
 RL Oncogene 19:4405-4416 (2000).  
 [6]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Colorectal carcinoma;  
 RA Lahm H., Siebert H.-C., Andre S., Hoefflich A., Diehl D., Sordat B.,  
 RA Kaltner H., Wolf E., Gabius H.-J.;  
 RT "Coca (Colorectal carcinoma-derived) galectin-8 variant I full-length  
 RT cDNA from a human colorectal carcinoma cell line.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 [7]  
 RP SEQUENCE FROM N.A.  
 RA Cobley V.;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 [8]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Brain, and Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,  
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1; Synonyms=1;  
 CC IsoId=O00214-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=O00214-2; Sequence=VSP\_003094;  
 CC Name=3;  
 CC IsoId=O00214-3; Sequence=VSP\_003095;  
 CC -!- TISSUE SPECIFICITY: Ubiquitous. Selective expression by prostate  
 CC carcinomas versus normal prostate and benign prostatic  
 CC hypertrophy.  
 CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding  
 CC domains.  
 CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; L78132; AAB51605.1; -;  
 DR EMBL; X91790; CAA62904.1; ALT\_INIT.  
 DR EMBL; AF074000; AAD45402.1; -;  
 DR EMBL; AF074001; AAD45403.1; -;  
 DR EMBL; AF074002; AAD45404.1; -;  
 DR EMBL; AF193806; AAF19370.1; ALT\_INIT.  
 DR EMBL; AF193805; AAF19370.1; JOINED.  
 DR EMBL; AF342815; AAK16735.1; ALT\_INIT.  
 DR EMBL; AL136105; CAC15946.1; ALT\_INIT.  
 DR EMBL; AL136105; CAC15947.1; ALT\_INIT.  
 DR EMBL; BC015818; AAL15818.1; -;  
 DR EMBL; BC016486; AAL16486.2; -;  
 DR HSSP; P17931; 1A3K.  
 DR Genew; HGNC:6569; LGAL8.  
 DR H-invDB; HIX0001724; -;  
 DR MIM; 606099; -;  
 DR GO; GO:0005635; C:extracellular space; TAS.  
 DR GO; GO:0005529; F:sugar binding; TAS.  
 DR InterPro; IPR008985; ConA like lec\_gl.  
 DR InterPro; IPR001079; Galectin.  
 DR Pfam; PF00337; Gal-bind lectin; 2.  
 DR PROSITE; PS00309; GALACTIN; 1.  
 DR Alternative splicing; Antigen; Galectin; Lectin; Polymorphism; Repeat.  
 FT DOMAIN 1 153  
 FT Linker.  
 FT DOMAIN 154 184  
 FT DOMAIN 185 316  
 FT SITE 248 254  
 FT VARSPPLIC 182 182  
 FT



```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Whiting J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanches A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RP STRAIN=FVB/N; TISSUE=Mammary tumor;
RC Strausberg R.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
CC EMBL; BC003754; AA03754.1; -.
DR HSSP; P17931; 1A3K.
DR MGI; 109496; Lgals9.
DR GO; GO:0005229; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALACTIN; 2.
KW Galectin; Lectin.
SQ SEQUENCE 322 AA; 36545 MW; 032D77400737562E CRC64;

Query Match 15.0%; Score 102; DB 2; Length 322;
Best Local Similarity 24.4%; Pred. No. 0.017;
Matches 32; Conservative 27; Mismatches 68; Indels 4; Gaps 2;

Qy 2 SPRLEVPCHALPQGLSPGQVIVRGLVLPQPKHTVSLR---DOAAHAPVTLRASFAADR 58
Db 187 TPATYTPPTIPNGLYPSKSMISGNVLPDPTFRHNLRCGGDIAPHLNFRFNENAVVR 246

Qy 59 TLAWISRWGQ-KKLISAPFLFPQRFVLLFQEGGLKALNGQGLGATSMNQALRQL 117
Db 247 NTQINNSWGQERSLLGRMPFSRQGSFVSWIICGHCFCFVAVNGQHMCEYVHRLKQLQDI 306

Qy 118 RELRIGSGVQL 128
Db 307 NTLVAGDIQL 317

RESULT 16
LEG9 MOUSE
ID _LEG9 MOUSE STANDARD; PRT; 353 AA.
AC O08573; O08572;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Galectin-9
GN Name=Lgals9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney, and Small intestine;
RX MEDLINE=97190351; PubMed=9038233; DOI=10.1074/jbc.272.9.6078;
RA Wada J., Kanwar Y.S.;
RT "Identification and characterization of galectin-9, a novel beta-
RT galactoside-binding mammalian lectin.";
RL J. Biol. Chem. 272:6078-6086(1997).
[2]
RN CHARACTERIZATION.

```

```

RX MEDLINE=97298141; PubMed=9153289;
RA Wada J., Ota K., Kumar A., Wallner E.I., Kanwar Y.S.;
RT "Developmental regulation, expression, and apoptotic potential of
RT galectin-9, a beta-galactoside binding lectin.";
RL J. Clin. Invest. 99:2452-2461(1997).
CC -!- FUNCTION: Binds galactosides. May play a role in thymocyte-
CC epithelial interactions relevant to the biology of the thymus.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. May also be secreted by a non-
CC classical secretory pathway.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=Long; IsoId=O08573-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O08573-2; Sequence=VSP_003097;
CC -!- TISSUE SPECIFICITY: Accentuated expression in liver and thymus of
CC embryo, detected in embryonic heart, brain, lung, liver, and
CC kidney. Highly expressed in adult thymus, small intestine, and
CC liver, and to a lesser extent in lung, kidney, spleen, cardiac,
CC and skeletal muscle. Barely detectable in brain and reticulocyte.
CC The long form is expressed exclusively in the small intestine.
CC -!- DEVELOPMENTAL STAGE: The expression increased with successive
CC stages of embryonic development.
CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding
CC domains.
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; U55061; AB51190.1; -.
DR EMBL; U55060; AB51189.1; -.
DR HSSP; P17931; 1A3K.
DR MGI; 109496; Lgals9.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR PROSITE; PS00309; GALACTIN; 2.
KW Alternative splicing; Galectin; Lectin; Repeat.
FT DOMAIN 1 147 Galactin 1.
FT DOMAIN 148 204 Linker.
FT DOMAIN 205 353 Galactin 2.
FT SITE 81 87 Beta-galactoside binding 1 (By
FT similarity).
FT SITE 285 291 Beta-galactoside binding 2 (By
FT similarity).
FT VARSPLIC 148 178 Missing (in isoform Short).
FT SEQUENCE 353 AA; 40036 MW; B54036F5E280C531 CRC64;
SQ SEQUENCE 353 AA; 40036 MW; B54036F5E280C531 CRC64;

Query Match 15.0%; Score 102; DB 1; Length 353;
Best Local Similarity 24.4%; Pred. No. 0.019;
Matches 32; Conservative 27; Mismatches 68; Indels 4; Gaps 2;

Qy 2 SPRLEVPCHALPQGLSPGQVIVRGLVLPQPKHTVSLR---DOAAHAPVTLRASFAADR 58
Db 218 TPATYTPPTIPNGLYPSKSMISGNVLPDPTFRHNLRCGGDIAPHLNFRFNENAVVR 277

Qy 59 TLAWISRWGQ-KKLISAPFLFPQRFVLLFQEGGLKALNGQGLGATSMNQALRQL 117
Db 278 NTQINNSWGQERSLLGRMPFSRQGSFVSWIICGHCFCFVAVNGQHMCEYVHRLKQLQDI 337

Qy 118 RELRIGSGVQL 128
Db 338 NTLVAGDIQL 348

```

```

RESULT 17
Q6Q2P2 PRELIMINARY; PRT; 323 AA.
AC Q6Q2P2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Galectin 9.
GN Name=UAT;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Bannasch D.L., Ryon J.R., Bannasch M.J., Schaible R.H., Breen M.,
RA Ling G.;
RT "Exclusion of galectin 9 as a candidate gene for hyperuricosuria in
RT the Dalmatian dog";
RL Anim. Genet. 35:326-328 (2004).
CC -1- SIMILARITY: Belongs to the galectin (galectin/S-lectin) family.
DR EMBL; AY521549; AA880311.1; -.
DR HSP; P17931; IA3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 2.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALAPTIN; 2.
KW Galectin; Lectin.
SQ SEQUENCE 323 AA; 36026 MW; 2C4FA3644A0EB64F CRC64;

Query Match 14.5%; Score 99; DB 2; Length 323;
Best Local Similarity 25.2%; Pred. No. 0.036;
Matches 33; Conservative 27; Mismatches 67; Indels 4; Gaps 2;

QY 2 SPRLEPCSHALPGQLSPGQVIVRGVLQEPKHTVSLR---DQAAHAPVTLRASPADR 58
DB 188 TPYVPMFFTSIPGGLYPSKSIIVSGTVLPCKARFHNLRSGNDIAFHLNPRFNENVVR 247
QY 59 TLAWISRWG-QKKLISAPFFYPORFPEVLLLFQEGGLKLALNGGLGATSMNQQALEQL 117
DB 248 NMQLNNSWGSEERSLPRKMPFVQGSFVIMCEGHCFKVAVDGEHLFEYHYRLKNLAI 307
QY 118 RELRISGSVOL 128
DB 308 NNMEVAGDVQL 318

RESULT 18
Q6P7Q6 PRELIMINARY; PRT; 322 AA.
AC Q6P7Q6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lactin, galactose binding, soluble 9.
GN Name=Lgals9;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Pituitary gland;
RA MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Pituitary gland;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the galectin (galectin/S-lectin) family.
DR EMBL; BC061566; AAH61566.1; -.
DR HSP; P17931; IA3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 2.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALAPTIN; 2.
KW Galectin; Lectin.
SQ SEQUENCE 322 AA; 36341 MW; CDD414A6FD1BA9DD CRC64;

Query Match 14.4%; Score 98; DB 2; Length 322;
Best Local Similarity 23.7%; Pred. No. 0.047;
Matches 31; Conservative 27; Mismatches 69; Indels 4; Gaps 2;

QY 2 SPRLEPCSHALPGQLSPGQVIVRGVLQEPKHTVSLR---DQAAHAPVTLRASPADR 58
DB 187 TPATYTPFFTSIPGGLYPSKSIIVSGTVLPCKARFHNLRSGNDIAFHLNPRFNENVVR 246
QY 59 TLAWISRWG-QKKLISAPFFYPORFPEVLLLFQEGGLKLALNGGLGATSMNQQALEQL 117
DB 247 NTQINNSWGSEERSLPRKMPFVQGSFVIMCEGHCFKVAVDGEHLFEYHYRLKNLPDI 306
QY 118 RELRISGSVOL 128
DB 307 NTLEVAGDIQL 317

RESULT 19
LEG4 MOUSE
ID LEG4 MOUSE STANDARD; PRT; 326 AA.
AC Q8K419; O88353; Q91X74;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Galectin-4 (lactose-binding lectin 4).
GN Name=Lgals4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/c; TISSUE=Liver;
RA Maly P., Jenikova G., Cummings R.D.;
RT "Molecular cloning and tissue distribution of mouse galectin-4.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6; TISSUE=Colon;
RA Hokama A., Tanaka Y., Mizoguchi A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

```



```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR ENBL; U59462; AAB51192.1; -
DR ENBL; U72741; AAB68592.1; -
DR ENBL; U67958; AAB48591.1; -
DR HSP; P17931; IA3K.
DR RGD; 3005; Lgal89.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR PROSITE; PS00309; GALACTIN; 2.
KW Alternative splicing; Galectin; Ion transport; Lectin; Repeat.
FT DOMAIN 1 147 Galectin 1.
FT DOMAIN 148 205 Linker.
FT DOMAIN 206 354 Galactin 2.
FT SITE 81 87 Beta-galactoside binding 1 (By
FT SITE 286 292 Beta-galactoside binding 2 (By
FT SITE 148 179 Missing (in isoform Short).
FT VARSPLIC 148 179 /FTid=VSP_003098.
FT SEQUENCE 354 AA; 39946 MW; 6574F960B2EAF37C CRC64;
Query Match 14.4%; Score 98; DB 1; Length 354;
Best Local Similarity 23.7%; Pred. No. 0.052;
Matches 31; Conservative 27; Mismatches 69; Indels 4; Gaps 2;
QY 2 SPRLVPCSHALQGLSPGVIIVRGLVLOEPKHTVSLR---DQAAHAPVTLRASPADR 58
DB 219 TPAYTIFFTSIPNGFYPSKINISNVGLPDKRFINLRGCGDIAPHNPRNEKVVVR 278
QY 59 TLAWISRWG-OKKLISAPFLFYPPQRFVEVLLFQGGKLGALNGQGLGATSMNQALEQL 117
DB 279 NTQINNSWGEERSLPGRMFNRGQSFSVILCEGHCFKAVDQGHICEYHYHRLKNLPDI 338
QY 118 RELRISGSVOL 128
DB 339 NTLVAGDIQL 349
RESULT 21
Q7TPX9 PRELIMINARY; PRT; 165 AA.
AC Q7TPX9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haleh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbalano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR ENBL; BC052838; AAB52838.1; -
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Galectin; Hypothetical protein; Lectin.
SQ SEQUENCE 165 AA; 18270 MW; 351BDDAF65B51D26 CRC64;
Query Match 14.3%; Score 97.5; DB 2; Length 165;
Best Local Similarity 24.6%; Pred. No. 0.024;
Matches 30; Conservative 29; Mismatches 54; Indels 9; Gaps 3;
QY 16 GLSPGVIIVRGLVLOEPKHTVSL-----RQAAHAPVTLRASPADRTL--AWISRWG 67
DB 40 GMPGKVLVGVINDNPKSFAISLTCGSEDDPADVAIEKLVVFNQVFRNSCISGES 99
QY 68 OKKLISAP-FLFYPPQRFVEVLLFQGGKLGALNGQGLGATSMNQALEQLRLRISGSV 126
DB 100 DEENLAPFPFPVDDPPFRMEIFCQPCFRLVDGHHLLFDYHRIQTLAISDKISGDL 159
QY 127 QL 128
DB 160 QI 161
RESULT 22
Q7ZTB8 PRELIMINARY; PRT; 308 AA.
AC Q7ZTB8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin family xgalectin-IIib.
GN Name=xgalectin-IIib;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22552404; PubMed=12538594; DOI=10.1074/jbc.M209008200;
RA Shoji H., Nishi N., Hiraehima M., Nakamura T.;
RT "Characterization of the Xenopus galectin family. Three structurally
RT different types as in mammals and regulated expression during
RT embryogenesis.";
RL J. Biol. Chem. 278:12285-12293(2003).
CC -1- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR ENBL; AB080017; BAC55883.1; -
DR HSP; P1116; ISLT.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR SMART; SM00276; GLECT; 2.
KW Galectin; Lectin.
SQ SEQUENCE 308 AA; 34808 MW; A1C22D04079048AE CRC64;
Query Match 14.2%; Score 97; DB 2; Length 308;
Best Local Similarity 24.6%; Pred. No. 0.057;

```



```
Matches 32; Conservative 26; Mismatches 60; Indels 12; Gaps 4;
Qy 10 SHALPQ-----GLSPGVIIIRGLVLQEPKHTVSLRDOAAHAPVTLRASFAADRTLAWI 63
Db 175 SYAMPYQNIYGLFPFKSTIVIRGTVPANPRPHNLKFXHGTA-LHNPFRFDETRVNR 233
Qy 64 SR-----WG-QKKLISAPFLFYQRFVFFVLLLFQEGGLKALNGQGLGATSMNQALBQLR 118
Db 234 SHLNGSWGNEERNLPRGMCFPAGFSFVIEIRCEQHAPKVVNGAQICEFHFRVHQFQID 293
Qy 119 ELRISGSVOL 128
Db 294 TLQIDGDVVL 303

RESULT 23
LEG6_MOUSE
ID LEG6_MOUSE STANDARD; PRT; 301 AA.
AC 054891; 088352;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galectin-6.
GN Name=Lgals6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=98112847; PubMed=9446608; DOI=10.1074/jbc.273.5.2954;
RA Gitt M.A., Colnot C., Poirier F., Nani K.J., Barondes S.H.,
RA Leffler H.;
RT "Galectin-4 and galectin-6 are two closely related lectins expressed
RT in mouse gastrointestinal tract.";
RL J. Biol. Chem. 273:2954-2960(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=98112848; PubMed=9446609; DOI=10.1074/jbc.273.5.2961;
RA Gitt M.A., Xia Y.-R., Atchison R.E., Lusis A.J., Barondes S.H.,
RA Leffler H.;
RT "Sequence, structure, and chromosomal mapping of the mouse Lgals6
RT gene encoding galectin-6.";
RL J. Biol. Chem. 273:2961-2970(1998).
CC -!- DOMAIN: Contains two homologous but distinct carbohydrate-binding
CC domains.
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF026799; AAC04508.1; -.
DR EMBL; AF026796; AAC04508.1; JOINED.
DR EMBL; AF026797; AAC04508.1; JOINED.
DR EMBL; AF026798; AAC04508.1; JOINED.
DR EMBL; AF026794; AAC27244.1; -.
DR HSSP; PI7931; 1A3K.
DR MGD; MGI:107535; Lgals6.
DR InterPro; IPR008985; ConA_like_lect_1.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 2.
DR SMART; SM00276; GLECT; 2.
DR PROSITE; PS00309; GALAPTIN; 2.
DR Galectin; Lactin; Repeat.
KW DOMAIN 1 151 Galectin 1.
FT DOMAIN 152 160 Linker.
```

```
FT DOMAIN 160 301 Galaptin 2.
SQ SEQUENCE 301 AA; 34112 MW; 9A4DD09944EDFAB9 CRC64;
Query Match 14.1%; Score 96; DB 1; Length 301;
Best Local Similarity 27.2%; Pred. No. 0.071;
Matches 37; Conservative 22; Mismatches 61; Indels 16; Gaps 3;
Qy 7 VPCSHALPQGLSPGVIIIRGLVLQEPKHTVSLR-----DQAAH-----APVTLRASFA 56
Db 171 LPVVGALQGGFTVRRTIIRKGYVLPKTAKEFAINFRVGSSEDIALHINPRIGDCLVRNSYM 230
Qy 57 DRTLAWISRWGQKKLISAPFLFYQRFVFFVLLLFQEGGLKALNGQGLGATSMNQALBQLR 116
Db 231 N-----GSMGTEERVMVAYNPFPGQGFDFLSIRCGMDRFKVFANGIHLNFESHFRQALRK 284
Qy 117 LRELISGSVOLYCVH 132
Db 285 INTLEINGDITLSYVH 300

RESULT 24
O8QGD9 PRELIMINARY; PRT; 332 AA.
ID O8QGD9;
AC O8QGD9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galectin-3TMI isoform containing transmembrane spanning domain.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22013973; PubMed=1186849; DOI=10.1074/jbc.M109578200;
RA Gorski J.P., Liu F.T., Artigues A., Castagna L.F., Osodoby P.;
RT "New alternatively spliced form of galectin-3, a member of the beta-
RT galactoside-binding animal lectin family, contains a predicted
RT transmembrane-spanning domain and a leucine zipper motif.";
RL J. Biol. Chem. 277:18840-18848(2002).
CC -!- SIMILARITY: Belongs to the galectin (galaptin/s-lectin) family.
DR EMBL; AF479564; AAJ91920.1; -.
DR HSSP; PI7931; 1A3K.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA_like_lect_1.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind_lectin; 1.
DR SMART; SM00276; GLECT; 1.
DR PROSITE; PS00309; GALAPTIN; UNKNOWN_1.
DR Galectin; Lactin; Transmembrane.
SQ SEQUENCE 332 AA; 35484 MW; CF3203BC99418CB1 CRC64;
Query Match 14.1%; Score 96; DB 2; Length 332;
Best Local Similarity 23.2%; Pred. No. 0.08;
Matches 33; Conservative 27; Mismatches 62; Indels 20; Gaps 3;
Qy 3 PRLEVPCSHALPQGLSPGVIIIRGLVLQEPKHTVSLR-----DQAAH-----APV 49
Db 193 PLQKVPYDPLPLPAGLMPRLITITGVNSNPNRFSLDKFGQDIAFHPNPRFKEDHKRVI 252
Qy 50 TLRASFAADRTLAWISRWGQKKLISAPFLFYQRFVFFVLLLFQEGGLKALNGQGLGATSM 109
Db 253 VCNMFQ-----NNWKEERTAPRFPFPFGTGFQKLVQCEGDHFKVAVNDHALHQLQNF 305
Qy 110 NQQALEQLRELISGSVOLYCV 131
Db 306 REKKLNEITKLCIAGDITLTSV 327

RESULT 25
```





```

DR EMBL; BC054324; AAH54324.1; -.
DR HSP; P17931; 1A3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLECT; 1.
KW Galectin; Lectin.
FT NON TER
SQ SEQUENCE 204 AA; 22637 MW; 82532249A1339BE1 CRC64;

Query Match 13.8%; Score 94; DB 2; Length 204;
Best Local Similarity 24.3%; Pred. No. 0.074;
Matches 37; Conservative 25; Mismatches 62; Indels 28; Gaps 6;

QY 1 MSPRLVPCSH-----ALPQGLSPQVIVRGLVLPQEPKHTVSLR-----DQAAH-- 46
DB 60 LPMWGPPIHLPFPKAMIPGGMIPKRTVIMKGLVNSNAKFIQISFKVGYTNDIALHIN 119
QY 47 -----APVTLRASPADRTLAWISRWG--QKKLISAPFLFYQRFQFVLLLFQEGGLKIALN 100
DB 120 PRLNKNLTIRNSFINGT-----WGEEDKDVKNP--FHQGEHFDISIRSGEKQYKVVN 171
QY 101 GQGLGATSMNQQAELRLRISGSVOLYCVH 132
DB 172 GYHCFNYPHRLTNLQQVDTLEADGDIKLCFVH 203

RESULT 29
Q7ZTB6 PRELIMINARY; PRT; 319 AA.
AC Q7ZTB6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Galectin family xgalectin-Via.
GN Name:xgalectin-Via;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22552404; PubMed=12538594; DOI=10.1074/jbc.M209008200;
RA Shoji H., Nishi N., Hirashima M., Nakamura T.;
RT "Characterization of the Xenopus galectin family. Three structurally
RT different types as in mammals and regulated expression during
RT embryogenesis.";
RL J. Biol. Chem. 278:12285-12293 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AB080019; BAC55885.1; -.
DR EMBL; BC077627; AAH77627.1; -.
DR HSP; P17931; 1A3K.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 2.
DR SMART; SM00276; GLECT; 2.
KW Galectin; Lectin.
SQ SEQUENCE 319 AA; 35369 MW; 2C12918A64B1BBFB CRC64;

Query Match 13.8%; Score 94; DB 2; Length 319;
Best Local Similarity 24.3%; Pred. No. 0.13;
Matches 37; Conservative 25; Mismatches 62; Indels 28; Gaps 6;

QY 1 MSPRLVPCSH-----ALPQGLSPQVIVRGLVLPQEPKHTVSLR-----DQAAH-- 46
DB 175 LPMWGPPIHLPFPKAMIPGGMIPKRTVIMKGLVNSNAKFIQISFKVGYTNDIALHIN 234
QY 47 -----APVTLRASPADRTLAWISRWG--QKKLISAPFLFYQRFQFVLLLFQEGGLKIALN 100
DB 235 PRLNKNLTIRNSFINGT-----WGEEDKDVKNP--FHQGEHFDISIRSGEKQYKVVN 286
QY 101 GQGLGATSMNQQAELRLRISGSVOLYCVH 132
DB 287 GYHCFNYPHRLTNLQQVDTLEADGDIKLCFVH 318

RESULT 30
Q8IGN9 PRELIMINARY; PRT; 218 AA.
AC Q8IGN9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative salivary galectin.
OS Anopheles stephensi (Indo-Pakistan malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=30069;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary glands;
RX MEDLINE=22710796; PubMed=12826099; DOI=10.1016/S0965-1748(03)00067-5;
RA Valenzuela J.G., Francischetti I.M.B., Pham V.M., Garfield M.K.,
RA Ribeiro J.M.C.;
RT "Exploring the salivary gland transcriptome and proteome of the
RT Anopheles stephensi mosquito.";
RL Insect Biochem. Mol. Biol. 33:717-732 (2003).
CC -!- SIMILARITY: Belongs to the galectin (galactin/S-lectin) family.
DR EMBL; AY162251; AAO06842.1; -.
DR HSP; P47929; 1BKZ.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001079; Galectin.
DR Pfam; PF00337; Gal-bind lectin; 1.
DR SMART; SM00276; GLECT; 1.

```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:36:15 ; Search time 172 Seconds

(without alignments)  
1315.435 Million cell updates/sec

Title: US-10-816-042-18

Perfect score: 3103

Sequence: 1 DAHKEVAHRFKDLGEENFK.....TCFAEEGKKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: Geneseqp16Dec04:\*

2: Geneseqp1980s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003ss:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3103	100.0	585	1 AAP90388	Aap90388 Mature hu
2	3103	100.0	585	2 AAR05318	Aar05318 Human ser
3	3103	100.0	585	2 AAR08457	Aar08457 Human ser
4	3103	100.0	585	2 AAR80301	Aar80301 Human ser
5	3103	100.0	585	2 AAO20111	AAO20111 HSA prote
6	3103	100.0	585	3 AAY84873	Aay84873 Amino aci
7	3103	100.0	585	3 AAY83946	Aay83946 Yeast cod
8	3103	100.0	585	4 AAM52567	Aam52567 Mature hu
9	3103	100.0	585	4 AAE12403	AAE12403 Human alb
10	3103	100.0	585	4 AAE13129	AAE13129 Human alb
11	3103	100.0	585	4 AAE13399	AAE13399 Human alb
12	3103	100.0	585	4 ABB79006	Abb79006 Human mat
13	3103	100.0	585	4 AAE08578	AAE08578 Human ser
14	3103	100.0	585	5 AAU75220	AAU75220 Mature fo
15	3103	100.0	585	5 ABJ00986	Abj00986 B lymphoc
16	3103	100.0	585	5 ABG63321	ABG63321 Human ser
17	3103	100.0	585	5 ABG33847	ABG33847 Human B L
18	3103	100.0	585	5 ABG71291	ABG71291 Glycosyla
19	3103	100.0	585	6 ABR55695	ABr55695 Human alb
20	3103	100.0	585	7 ABR42606	ABr42606 Human ser
21	3103	100.0	585	7 ADC16767	Adc16767 Human ser
22	3103	100.0	585	7 ADD06469	Add06469 Human ser
23	3103	100.0	585	7 ADD68016	Add68016 Mature fo
24	3103	100.0	585	7 ADF15951	Adf15951 Human alb
25	3103	100.0	585	7 ADH21530	Adh21530 Mature hu

26	3103	100.0	585	7 ADK40322	Adk40322 Mature hu
27	3103	100.0	585	8 ABM79785	Abm79785 Human ser
28	3103	100.0	585	8 ADJ50489	Adj50489 Human ser
29	3103	100.0	585	8 ADJ64278	Adj64278 Human alb
30	3103	100.0	585	8 ADL76536	Adl76536 Albumin f
31	3103	100.0	585	8 ADP82617	Adp82617 Human alb
32	3103	100.0	585	8 ADP82618	Adp82618 Human ace
33	3103	100.0	585	8 ADR88124	Adr88124 Human alb
34	3103	100.0	609	3 AAB36542	Aab36542 Recombina
35	3103	100.0	609	3 AAY78147	Aay78147 Pre human
36	3103	100.0	609	3 AAB36549	Aab36549 Recombina
37	3103	100.0	609	6 ABUS7252	Abus7252 Human ser
38	3103	100.0	609	6 ABUS7253	Abus7253 Human ser
39	3103	100.0	609	7 ADD06471	Add06471 Human ser
40	3103	100.0	609	7 ADH21582	Adh21582 Human alb
41	3103	100.0	609	7 ADH21582	Adh21582 Human ser
42	3103	100.0	609	7 ADH21582	Adh21582 Human ser
43	3103	100.0	609	8 ADN16134	Adn16134 Protein s
44	3103	100.0	609	8 ADE77205	Ade77205 Human pro
45	3103	100.0	609	8 ADL13248	Adl13248 Human ste
46	3103	100.0	609	8 ADP55998	Adp55998 Human PRO
47	3103	100.0	610	2 AAR39510	Aar39510 Chimeric
48	3103	100.0	616	6 AAE30916	Aae30916 Val8-GLP-
49	3103	100.0	619	7 ADF14974	Adf14974 Human alb
50	3103	100.0	619	7 ADF14973	Adf14973 Human alb
51	3103	100.0	619	7 ADF16250	Adf16250 Human alb
52	3103	100.0	621	7 ADF16247	Adf16247 Human alb
53	3103	100.0	623	7 ADF16461	Adf16461 Human alb
54	3103	100.0	623	7 ADF16466	Adf16466 Human alb
55	3103	100.0	623	7 ADF16492	Adf16492 Human alb
56	3103	100.0	623	7 ADH21791	Adh21791 Human alb
57	3103	100.0	623	7 ADH21776	Adh21776 Human alb
58	3103	100.0	623	7 ADH21780	Adh21780 Human alb
59	3103	100.0	624	6 AAE30919	Aae30919 Human ser
60	3103	100.0	631	6 AAE30917	Aae30917 Val8-GLP-
61	3103	100.0	634	7 ADF16439	Adf16439 Human alb
62	3103	100.0	634	7 ADF16441	Adf16441 Human alb
63	3103	100.0	635	7 ADF16547	Adf16547 Human alb
64	3103	100.0	635	7 ADF15012	Adf15012 Human alb
65	3103	100.0	635	7 ADF15011	Adf15011 Human alb
66	3103	100.0	635	7 ADF15005	Adf15005 Human alb
67	3103	100.0	635	7 ADF15013	Adf15013 Human alb
68	3103	100.0	635	7 ADH21296	Adh21296 Human alb
69	3103	100.0	635	7 ADH21297	Adh21297 Human alb
70	3103	100.0	635	7 ADH21294	Adh21294 Human alb
71	3103	100.0	635	7 ADH21295	Adh21295 Human alb
72	3103	100.0	636	7 ADF16440	Adf16440 Human alb
73	3103	100.0	636	7 ADF16463	Adf16463 Human alb
74	3103	100.0	636	7 ADF16442	Adf16442 Human alb
75	3103	100.0	636	7 ADF16497	Adf16497 Human alb
76	3103	100.0	637	7 ADF16468	Adf16468 Human alb
77	3103	100.0	637	7 ADF16467	Adf16467 Human alb
78	3103	100.0	637	7 ADH21781	Adh21781 Human alb
79	3103	100.0	637	7 ADH21782	Adh21782 Human alb
80	3103	100.0	638	7 ADF16203	Adf16203 Human alb
81	3103	100.0	638	7 ADF42052	Adf42052 Plasmid p
82	3103	100.0	638	8 ADR90044	Adr90044 pDB2300X1
83	3103	100.0	639	7 ADF15120	Adf15120 Human alb
84	3103	100.0	639	7 ADF16198	Adf16198 Human alb
85	3103	100.0	639	7 ADF15119	Adf15119 Human alb
86	3103	100.0	639	7 ADF16208	Adf16208 Human alb
87	3103	100.0	639	7 ADF15118	Adf15118 Human alb
88	3103	100.0	639	7 ADF15116	Adf15116 Human alb
89	3103	100.0	639	7 ADH21332	Adh21332 Human alb
90	3103	100.0	639	7 ADH21335	Adh21335 Human alb
91	3103	100.0	639	7 ADH21333	Adh21333 Human alb
92	3103	100.0	639	7 ADH21334	Adh21334 Human alb
93	3103	100.0	640	6 AAE30920	Aae30920 Exendin-4
94	3103	100.0	640	6 AAE30918	Aae30918 Gly8-Glu2
95	3103	100.0	640	7 ADF14992	Adf14992 Human alb
96	3103	100.0	640	7 ADF16505	Adf16505 Human alb
97	3103	100.0	640	7 ADF16506	Adf16506 Human alb
98	3103	100.0	640	7 ADF14984	Adf14984 Human alb

99	3103	100.0	640	7	AdF16530	Human alb
100	3103	100.0	640	7	AdH21818	Human alb
ALIGNMENTS						
RESULT 1						
AAP90388						
ID	AAP90388 standard; protein; 585 AA.					
AC	AAP90388;					
XX						
DT	24-OCT-2003 (revised)					
DT	25-MAR-2003 (revised)					
DT	01-NOV-1989 (first entry)					
XX						
DE	Mature human serum albumin polypeptide.					
XX						
KW	Human serum albumin; mature protein; new polypeptides; plasma expanders.					
XX						
OS	Homo sapiens; (Human).					
XX						
PN	EP322094-A.					
XX						
PD	28-JUN-1989.					
XX						
PF	25-OCT-1988; 88EP-00310000.					
XX						
PR	30-OCT-1987; 87GB-00025529.					
XX						
PA	(DELZ ) DELTA BIOTECHNOLOGY LTD.					
XX						
PI	Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;					
XX						
DR	WPI; 1989-186464/26.					
DR	N-PSDB; AAN90128.					
XX						
PT	New N-terminal fragments of human serum albumin - esp. useful as blood					
PT	plasma expanders.					
XX						
PS	Disclosure; Fig 2; 20pp; English.					
XX						
CC	Mature protein of human serum albumin (see corresp. AAN90128). Used to					
CC	make new N-terminal fragments which are used as plasma expanders, or as					
CC	substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-					
CC	2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS					
CC	field)					
XX						
SQ	Sequence 585 AA;					
Query Match 100.0%; Score 3103; DB 1; Length 585;						
Best Local Similarity 100.0%; Pred. No. 1.9e-254;						
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	DAHKSEVAHRFKDGLGENFKALVLI AFAYLQQCPFDHVKLVNVEVTEFAKTCVADESAAE 60				
DB	1	DAHKSEVAHRFKDGLGENFKALVLI AFAYLQQCPFDHVKLVNVEVTEFAKTCVADESAAE 60				
QY	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120				
DB	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120				
QY	121	DVMCTAFHNDNEFTFLKKLYIYEIARRHPYFAPPELLFFAKRYKAAATECCQAADKAAACLLP 180				
DB	121	DVMCTAFHNDNEFTFLKKLYIYEIARRHPYFAPPELLFFAKRYKAAATECCQAADKAAACLLP 180				
QY	181	KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFPKAEFAEVSCLVTDLTG 240				
DB	181	KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFPKAEFAEVSCLVTDLTG 240				
QY	241	VHTECCHGDLLECCADRADLAKYICENQDSISSKLEKCECKPILLEKSHCIAEVENDEMPA 300				
DB	241	VHTECCHGDLLECCADRADLAKYICENQDSISSKLEKCECKPILLEKSHCIAEVENDEMPA 300				

Db	241	VHTECCHGDLLECCADRADLAKYICENQDSISSKLEKCECKPILLEKSHCIAEVENDEMPA	300
QY	301	DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLEKC	360
DB	301	DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLEKC	360
QY	361	CAAADPHECYAKVDFEFKPLVBEPPQNLIKQNCELFEQLGEYKFNALLVRYTKKVPQVST	420
DB	361	CAAADPHECYAKVDFEFKPLVBEPPQNLIKQNCELFEQLGEYKFNALLVRYTKKVPQVST	420
QY	421	PTLVEVSRNLGVGSKCKKHPEAKMPCAEEDYLSVVLNOLCVLHEKTPVSDRVTCKCTES	480
DB	421	PTLVEVSRNLGVGSKCKKHPEAKMPCAEEDYLSVVLNOLCVLHEKTPVSDRVTCKCTES	480
QY	481	LNNRRCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHKPRAT	540
DB	481	LNNRRCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHKPRAT	540
QY	541	KEQLKAVMDDFAAFVEKCKCCKADDKETCFABEGKKLVAAASQAALGL	585
DB	541	KEQLKAVMDDFAAFVEKCKCCKADDKETCFABEGKKLVAAASQAALGL	585
RESULT 2			
AAR05318			
ID	AAR05318 standard; protein; 585 AA.		
AC	AAR05318;		
XX			
DT	08-OCT-1990 (first entry)		
XX	Human serum albumin gene product.		
XX	Human serum albumin; HSA-A; yeast; ds.		
XX	Homo sapiens.		
OS	JP02117384-A.		
PN	01-MAY-1990.		
PD			
XX	26-OCT-1988; 88JP-00268302.		
PF			
XX	26-OCT-1988; 88JP-00268302.		
PR	(TOFU ) TOA NENRYO KOGYO KK.		
PA	WPI; 1990-176228/23.		
XX	N-PSDB; AAQ04719.		
DR	Human serum albumin prepn. by yeast host - by culturing transformed		
XX	plasmid yeast to produce serum, and removing it.		
PT	Disclosure; Page ?; -pp; Japanese.		
PS	Mature HSA-A may be produced using the sequence incorporated into a		
XX	plasmid vector with suitable controllers, and transferred to a yeast		
CC	expression system		
CC	Sequence 585 AA;		
XX	Query Match 100.0%; Score 3103; DB 2; Length 585;		
SQ	Best Local Similarity 100.0%; Pred. No. 1.9e-254;		
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	DAHKSEVAHRFKDGLGENFKALVLI AFAYLQQCPFDHVKLVNVEVTEFAKTCVADESAAE 60	
DB	1	DAHKSEVAHRFKDGLGENFKALVLI AFAYLQQCPFDHVKLVNVEVTEFAKTCVADESAAE 60	
QY	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120	
DB	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120	

RESULT 2

AAR05318

ID AAR05318 standard; protein; 585 AA.

XX AAR05318;

AC AAR05318;

XX 08-OCT-1990 (first entry)

DT Human serum albumin gene product.

DE Human serum albumin; HSA-A; yeast; ds.

XX Homo sapiens.

OS JP02117384-A.

PN 01-MAY-1990.

XX 26-OCT-1988; 88JP-00268302.

XX 26-OCT-1988; 88JP-00268302.

XX (TOFU ) TOA NENRYO KOGYO KK.

XX WPI; 1990-176228/23.

XX N-PSDB; AAQ04719.

XX Human serum albumin prepn. by yeast host - by culturing transformed

XX plasmid yeast to produce serum, and removing it.

XX Disclosure; Page ?; -pp; Japanese.

XX Mature HSA-A may be produced using the sequence incorporated into a

XX plasmid vector with suitable controllers, and transferred to a yeast

XX expression system

XX Sequence 585 AA;

Query Match

Best Local Similarity 100.0%; Score 3103; DB 2; Length 585;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGKEENFKALVLI AFAYLQQCPFDHVKLVNVEVTEFAKTCVADESAAE 60

DB 1 DAHKSEVAHRFKDGLGKEENFKALVLI AFAYLQQCPFDHVKLVNVEVTEFAKTCVADESAAE 60

QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120

DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120



QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db |||||  
 QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db |||||  
 QY 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
 Db |||||  
 QY 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
 Db |||||  
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
 Db |||||  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360  
 Db |||||  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360  
 Db |||||  
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
 Db |||||  
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
 Db |||||  
 QY 421 PTLVEVSRNLGKVGSKCKKHPBAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
 Db |||||  
 QY 421 PTLVEVSRNLGKVGSKCKKHPBAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
 Db |||||  
 QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
 Db |||||  
 QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
 Db |||||  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
 Db |||||  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
 Db |||||

## RESULT 3

AAR08457  
 ID AAR08457 standard; protein; 585 AA.

XX AAR08457;  
 AC AAR08457;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 16-APR-1991 (first entry)  
 XX  
 XX Human serum albumin.  
 XX  
 KW HSA; folding; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..303  
 FT /label= B  
 FT Region 123..585  
 FT /label= C  
 FT Region 123..303  
 FT /label= A  
 XX  
 JP02227079-A.  
 PN  
 PD 10-SEP-1990.  
 XX  
 PF 25-AUG-1989; 89JP-00217540.  
 XX  
 PR 06-OCT-1988; 88JP-00250926.  
 XX  
 XX (TOFU ) TONEN CORP.  
 XX  
 PA WPI; 1990-317325/42.  
 DR N-PSDB; AAQ06099.  
 XX  
 XX New human serum albumin fragments - used to bond to medicines and for  
 PT stable folding of protein(s).  
 XX  
 PS Claim 1; Fig 8; 24pp; Japanese.

XX Fragments A-C of HSA are expressed as fusion proteins with the signal  
 CC peptide of E. coli alkaline phosphatase. The fragments are selected for  
 CC their specific properties. The C-terminal truncated fragment, B, does not  
 CC bind long-chain fatty acids but does bind to various medicines at the  
 CC central region. The N-terminal truncated fragment, C, has good stability  
 CC in protein folding. The central segment, A, has characteristics of both B  
 CC and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD  
 CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-  
 CC 2003 to correct PR field.)  
 XX

SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-294;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAB 60  
 Db |||||  
 QY 1 DAHKEVAHRPKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAB 60  
 Db |||||  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
 Db |||||  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
 Db |||||  
 QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db |||||  
 QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db |||||  
 QY 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
 Db |||||  
 QY 181 KLDELDRDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
 Db |||||  
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
 Db |||||  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360  
 Db |||||  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLLRLAKTYETTLK 360  
 Db |||||  
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
 Db |||||  
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
 Db |||||  
 QY 421 PTLVEVSRNLGKVGSKCKKHPBAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
 Db |||||  
 QY 421 PTLVEVSRNLGKVGSKCKKHPBAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
 Db |||||  
 QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
 Db |||||  
 QY 481 LVNRRPCFSALEVDYTPVKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
 Db |||||  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
 Db |||||  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
 Db |||||

## RESULT 4

AAR0301  
 ID AAR0301 standard; protein; 585 AA.

XX AAR0301;  
 AC AAR0301;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 17-JAN-1996 (first entry)  
 XX  
 XX Human serum albumin.  
 XX  
 KW Serum albumin; HSA; aspartyl protease-3; Yap3p; Saccharomyces cerevisiae.  
 XX  
 OS Homo sapiens.  
 XX

PN W09523857-A1.  
XX  
PD 08-SEP-1995.  
XX  
PF 01-MAR-1995; 95WO-GB000434.  
XX  
PR 05-MAR-1994; 94GB-00004270.  
XX  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
PI Kerrywilliams SM, Gilbert SC;  
XX  
XX WPI; 1995-320572/41.  
DR N-PSDB; AAQ98695.  
XX  
XX Yeast with reduced levels of aspartyl protease 3 proteolytic activity -  
PT used to secrete human albumin without prodn. of the 45 kD fragment.  
XX  
XX Example 1; Page 26-28; 50pp; English.  
XX  
CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected  
CC to site-directed mutagenesis to investigate the role of endoproteases in  
CC the generation of a 45 kDa albumin fragment obtd. when the cDNA is  
CC expressed in S. cerevisiae. Mutations were: R410A, L408V, V409A;  
CC and R410A, K413Q, K414Q. The latter set of mutations, especially,  
CC improved stability of HSA to yeast Yap3p proteolytic cleavage, allowing  
CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct  
CC FI field.)  
XX  
SQ Sequence 585 AA;  
  
Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRIVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRIVRPEV 120  
  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
  
QY 181 KLDELREDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEYVSKLVTDLTK 240  
DB 181 KLDELREDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEYVSKLVTDLTK 240  
  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVVLRLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVVLRLRLAKTYETTTLEKC 360  
  
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGEYKFNQALLVRYTKKVPQVST 420  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGEYKFNQALLVRYTKKVPQVST 420  
  
QY 421 PTLVEVSNLGVSKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTTES 480  
DB 421 PTLVEVSNLGVSKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTTES 480  
  
QY 481 LVNRPFCFSALEVDSTYVPKFNARTFTFHADICTLSEKERQIKQTALVELVHKHPKAT 540  
DB 481 LVNRPFCFSALEVDSTYVPKFNARTFTFHADICTLSEKERQIKQTALVELVHKHPKAT 540  
  
QY 541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585

DB 541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585  
  
RESULT 5  
AAO20111  
ID AAO20111 standard; protein; 585 AA.  
XX  
XX AAO20111;  
AC AAO20111;  
XX  
DT 06-AUG-2002 (first entry)  
XX  
DE HSA protein sequence related to the growth hormone protein.  
XX  
KW Serum albumin-growth hormone fusion protein; growth hormone;  
KW Down's syndrome.  
XX  
OS Unidentified.  
XX  
PN KR99076789-A.  
XX  
PD 15-OCT-1999.  
XX  
PF 25-JUN-1998; 98KR-00704914.  
XX  
PR 30-DEC-1995; 95GB-00026733.  
PR 19-DEC-1996; 96WO-GB003164.  
XX  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
PI Ballance DJ;  
XX  
DR WPI; 1997-363680/33.  
DR N-PSDB; AAK99566.  
XX  
PT Serum albumin-growth hormone fusion protein - useful to treat growth  
PT hormone related diseases, e.g. Down's syndrome.  
XX  
PS Disclosure; Fig 6; 21pp; Korean.  
XX  
CC The invention relates to a serum albumin-growth hormone fusion protein -  
CC useful to treat growth hormone related diseases such as Down's syndrome.  
CC This sequence represents a HSA protein related to the serum albumin-  
CC growth hormone protein of the invention  
XX  
SQ Sequence 585 AA;  
  
Query Match 100.0%; Score 3103; DB 2; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRIVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRIVRPEV 120  
  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
  
QY 181 KLDELREDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEYVSKLVTDLTK 240  
DB 181 KLDELREDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEYVSKLVTDLTK 240  
  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCPEPLLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVVLRLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEYARRHPDYSVVLRLRLAKTYETTTLEKC 360

QY 361 CAADPHCYAKVDFBPKLVBEPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCYAKVDFBPKLVBEPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
QY 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585  
DB 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 6  
AAV84873  
ID AAY84873 standard; protein; 585 AA.  
XX  
AC AAY84873;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE Amino acid sequence of a human albumin protein.  
XX  
KW Human; albumin; ischemic state; serum protein; metal ion salt;  
KW peroperative ischemia; ischemia; myocardial infarction;  
KW progressive coronary artery disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "optionally acetylated, and claimed under claim  
FT 56"  
XX  
PN WO200020840-A1.  
XX  
PD 13-APR-2000.  
XX  
PF 01-OCT-1999; 99WO-US022905.  
XX  
PR 02-OCT-1998; 98US-00165581.  
PR 02-OCT-1998; 98US-00165926.  
PR 02-OCT-1998; 98US-0102738P.  
PR 11-JAN-1999; 99US-0115392P.  
XX  
PA (ISCH-) ISCHEMIA TECHNOLOGIES INC.  
XX  
PI Bar-Or D, Lau E, Winkler JV;  
XX  
DR WPI; 2000-303843/26.  
XX  
PT New method for the continuous detection of ischemic states comprises  
PT detecting and quantifying the existence of an alteration of the serum  
PT protein albumin.  
XX  
PS Disclosure; Page 97-100; 105pp; English.  
XX

XX The present sequence represents human albumin protein. The specification  
CC describes a method for the continuous detection of ischemic states. The  
CC method comprises detecting and quantifying the existence of an alteration  
CC of the serum protein albumin. The method comprises contacting a  
CC biological sample containing albumin from the patient with an excess  
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus  
CC of naturally occurring human albumin, to form a mixture containing bound  
CC metal ions and unbound metal ions, and then determining the amount of  
CC metal ions bound to the albumin N-terminus. The amount of bound metal  
CC ions is correlated to a known value to determine the occurrence or non-  
CC occurrence of an ischemic event. The methods are useful for detection of

CC ischemic states. The methods are also useful for distinguishing  
CC peroperative ischemia from ischemia caused by , amongst other things,  
CC myocardial infarctions and progressive coronary artery disease  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAE 60  
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVLRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVLRPEV 120  
QY 121 DVMCTAFHNDNETFLKKYLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMCTAFHNDNETFLKKYLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLUKCECEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLUKCECEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDKVCKNYAEAKVFLGMFLYEYARRHPDYSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 301 DLPSLAADFVSKDKVCKNYAEAKVFLGMFLYEYARRHPDYSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 361 CAADPHCYAKVDFBPKLVBEPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCYAKVDFBPKLVBEPQNLIKONCELFQQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540  
QY 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585  
DB 541 KEQLKAVNMDDFAAVFEKCKCKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 7  
AAV83946  
ID AAY83946 standard; protein; 585 AA.  
XX  
AC AAY83946;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Yeast codon-biased recombinant human serum albumin protein.  
XX  
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX  
OS Homo sapiens.  
OS Synthetic.  
PN CN1239103-A.  
XX  
PD 22-DEC-1999.  
XX  
PF 17-JUN-1998; 98CN-00102506.  
XX

```
PR 17-JUN-1998; 98CN-00102506.
XX (HAUJ-) HALJI BIOENGINEERING CO LTD.
XX LI S, Lu D;
XX WPI; 2000-351198/31.
DR N-PSDB; AAA10091.
XX
XX Process for preparing recombined human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX Disclosure; Fig 1; 44pp; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the complete sequence of the HSA encoded by the human gene with a yeast
CC codon bias. The invention also covers a recombinant expression vector,
CC yeast host cells carrying the recombinant expression vector and the
CC process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 3103; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.9e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
QY 121 DVMCTAFHDNEBTFLLKKLYEYIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNEBTFLLKKLYEYIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELRDSEKASSAKORLKASQKFGRAFKAFAVARLSQRPKAEFAEYVKLVTDLT 240
DB 181 KLDELRDSEKASSAKORLKASQKFGRAFKAFAVARLSQRPKAEFAEYVKLVTDLT 240
QY 241 VHTCECHGDLLECCADRDADLAKYICENODSISSSKLECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCECHGDLLECCADRDADLAKYICENODSISSSKLECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDYCKNYAEAKDVLGFLYVEYARRHPDYSVLLLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDYCKNYAEAKDVLGFLYVEYARRHPDYSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHCEYAKVDFEPFLVEEPQNLIKONCELFEOQLGEYKFQNALIVRVTKKVPQYST 420
DB 361 CAAADPHCEYAKVDFEPFLVEEPQNLIKONCELFEOQLGEYKFQNALIVRVTKKVPQYST 420
QY 421 PTLVEVSRLNLGVSKCKGHPKAPCAEDYLSVVLNQLCVLHBEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNLGVSKCKGHPKAPCAEDYLSVVLNQLCVLHBEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVDEVTYPKFNASTFTFHADICTLSKEKROIKKQATLVELVHKHPKAT 540
DB 481 LVNRRPCFSALVDEVTYPKFNASTFTFHADICTLSKEKROIKKQATLVELVHKHPKAT 540
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
```

```
RESULT 8
AAM52567
ID AAM52567 standard; protein; 585 AA.
XX
XX AAM52567;
XX
XX 05-FEB-2002 (first entry)
XX
XX Mature human serum albumin.
XX
XX Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
XX nootropic; neuroprotective; gene therapy; immune disorder; wound healing;
XX hyperproliferative disorder; renal disorder; cardiovascular disorder;
XX respiratory disorder; neurological disease; endocrine disorder;
XX reproductive system disorder; infectious disease;
XX gastrointestinal disorder.
XX
XX Homo sapiens.
XX
XX WO200179444-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012013.
XX
XX 12-APR-2000; 2000US-0229358P.
XX
XX 25-APR-2000; 2000US-0199384P.
XX
XX 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2001-616755/71.
XX
XX N-PSDB; ABA03057.
XX
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
XX useful in the treating immune system disorders (e.g. transplant
XX rejection), blood related disorders (e.g. myocardial infarction) and
XX hyperproliferative disorders.
XX
XX Claim 1; Fig 15; 606pp; English.
XX
XX The present invention relates to albumin fusion proteins, which comprise
XX a therapeutic protein and albumin. The present sequence is the protein
XX sequence for mature human serum albumin (HA), which was used to generate
XX the fusion proteins of the present invention. The albumin fusion proteins
XX are useful in the treatment, prevention, diagnosis, and/or detection of
XX diseases/disorders such as immune system disorders (e.g. transplant
XX rejection), blood related disorders (e.g. myocardial infarction),
XX hyperproliferative disorders (e.g. childhood acute myeloid leukemia),
XX renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
XX arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
XX neurological diseases (e.g. Alzheimer's disease), endocrine disorders
XX (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
XX infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
XX irritable bowel syndrome) and wound healing
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.9e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
QY 121 DVMCTAFHDNEBTFLLKKLYEYIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
```

Db 121 DWCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCECQAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240  
Qy 241 VHTCECHGDLLECADDRLAKYICENQDSISSKLUKCECKEPLLEKSHCIAEVNDMPA 300  
Db 241 VHTCECHGDLLECADDRLAKYICENQDSISSKLUKCECKEPLLEKSHCIAEVNDMPA 300  
Qy 301 DLPSLAADVESKDVCKNVAEAKDVLGNFLVEYARRHPDYSVLLRLRAKTYETTLK 360  
Db 301 DLPSLAADVESKDVCKNVAEAKDVLGNFLVEYARRHPDYSVLLRLRAKTYETTLK 360  
Qy 361 CAADPHECYAKVDFEFLVPEPQNLIKONCELFQELGEYKFQNALLVRYTKVPQVST 420  
Db 361 CAADPHECYAKVDFEFLVPEPQNLIKONCELFQELGEYKFQNALLVRYTKVPQVST 420  
Qy 421 PTLVEVRNLGVGSKCKCHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Db 421 PTLVEVRNLGVGSKCKCHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Qy 481 LVNRRPCFSALVEDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALVEDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVGHKPKAT 540  
Qy 541 KEQLKAVMDDDFAAFVEKCKADDDKCTCFAEKGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKADDDKCTCFAEKGKLVAAASQAALGL 585

RESULT 9  
AAE12403  
ID AAE12403 standard; protein; 585 AA.  
AC AAE12403;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human albumin (HA).  
XX  
KW Human; albumin; HA; immune system disorder; transplant rejection;  
KW blood related disorder; myocardial infarction; glomerulonephritis;  
KW hyperproliferative disorder; childhood acute myeloid leukaemia;  
KW renal cell carcinoma; cardiovascular disorder; vulnery; melanoma;  
KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukaemic;  
KW neurological disease; Alzheimer's disease; endocrine disorder; measles;  
KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;  
KW infectious disease; gastrointestinal disorder; wound healing; nontropic;  
KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;  
KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;  
KW antiarthritic; antirheumatic; renal disorder; antimicrobial.  
XX  
OS Homo sapiens.  
XX

FT Key Location/Qualifiers  
FT Domain /label= Loop\_I  
FT Domain /label= Loop\_II  
FT Domain /label= Loop\_III  
FT Domain /label= Loop\_IV  
FT Domain /label= Loop\_V  
FT Domain /label= Loop\_VI  
FT Domain /label= Loop\_VII  
FT Domain /label= Loop\_VIII

FT Domain 439..447  
FT /label= Loop\_IX  
FT Domain 461..475  
FT /label= Loop\_X  
FT Domain 478..486  
FT /label= Loop\_XI  
FT Domain 560..566  
FT /label= Loop\_XII  
XX  
PN WO200179480-A1.  
XX  
PD 25-OCT-2001.  
XX  
XX  
PF 12-APR-2001; 2001WO-US011991.  
XX  
PR 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
PR 21-DEC-2000; 2000US-0256931P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
DR WPI: 2001-616756/71.  
DR N-PSDB; AAD20005.  
XX  
PT Albumin fusion proteins comprising a therapeutic protein and albumin,  
PT useful in the treating metastatic renal cell carcinoma, metastatic  
PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human  
PT immunodeficiency virus) or infection.  
XX  
PS Claim 1; Fig 9; 394pp; English.  
XX  
CC The invention relates to human albumin (HA) fusion proteins and their  
CC corresponding nucleic acid sequences. Therapeutic proteins fused to  
CC albumin or its fragments have an extended shelf-life. The albumin fusion  
CC proteins are useful in the treatment, prevention, diagnosis, and/or  
CC detection of diseases, disorders such as immune system disorders (e.g.  
CC transplant rejection), blood related disorders (e.g. myocardial  
CC infarction), hyperproliferative disorders (e.g. childhood acute myeloid  
CC leukaemia, metastatic renal cell carcinoma, metastatic melanoma,  
CC malignant melanoma, renal cell carcinoma), renal disorders (e.g.  
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),  
CC respiratory disorders (e.g. non-allergic rhinitis), neurological diseases  
CC (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),  
CC reproductive system disorders (e.g. syphilis), infectious diseases (e.g.  
CC measles), gastrointestinal disorders (e.g. irritable bowel syndrome), HIV  
CC (human immunodeficiency virus) infection and wound healing. Nucleic acids  
CC encoding albumin fusion protein is used in gene therapy. The present  
CC sequence is human albumin  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254; Mismatches 0; Indels 0; Gaps 0;  
Matches 585; Conservative 0;  
Qy 1 DAHKSEVAHRFKDLGEENFKALVLIFAQYLOQCPEFDHVKLVNVEVTEFAKTCVADESAE 60  
Db 1 DAHKSEVAHRFKDLGEENFKALVLIFAQYLOQCPEFDHVKLVNVEVTEFAKTCVADESAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVRPEV 120  
Qy 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCECQAADKAACLLP 180  
Db 121 DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCECQAADKAACLLP 180  
Qy 181 KLDELDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240  
Db 181 KLDELDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240

QY	241	VHTECHGDLLECADDRADLAKYICENQDSSISSKLKECEKPLLEKSHCIAEVENDEMPA	300
Db	241	VHTECHGDLLECADDRADLAKYICENQDSSISSKLKECEKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRSLAKTYETTTLEKC	360
Db	301	DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYFYARRHPDYSVLLLRSLAKTYETTTLEKC	360
QY	361	CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFQQLGEYKFNQALLVRYTKKVPQVST	420
Db	361	CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFQQLGEYKFNQALLVRYTKKVPQVST	420
QY	421	PTLVEVSNLGVSKCKCHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES	480
Db	421	PTLVEVSNLGVSKCKCHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES	480
QY	481	LVNRRPCFSALEVDYTYVPKCFNAETFTTHADICTLSBKERQIKQTALVELVKHKPKAT	540
Db	481	LVNRRPCFSALEVDYTYVPKCFNAETFTTHADICTLSBKERQIKQTALVELVKHKPKAT	540
QY	541	KEQLKAVMDDFAAFVEKCKKADDDKTCFAESGKKLVAASQAALGL	585
Db	541	KEQLKAVMDDFAAFVEKCKKADDDKTCFAESGKKLVAASQAALGL	585
RESULT 10			
AAE13129			
ID	AAE13129	standard; protein; 585 AA.	
AC	AAE13129;		
XX	28-JAN-2002	(first entry)	
XX	Human albumin (HA).		
XX	Human; albumin; HA; fusion protein; therapeutic protein; vulnery;		
KW	immune system disorder; transplant rejection; blood related disorder;		
KW	myocardial infarction; hyperproliferative disorder; glomerulonephritis;		
KW	childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;		
KW	respiratory disorder; gene therapy; non-allergic rhinitis; neutropic;		
KW	neurological disease; Alzheimer's disease; reproductive system disorder;		
KW	endocrine disorder; pheochromocytoma; infectious disease; antiarthritic;		
KW	measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;		
KW	wound healing; antiinflammatory; immunosuppressive; neuroprotective;		
KW	cardiant; cytostatic; antileukaemic; antirheumatic; antimicrobial;		
KW	renal disorder.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Domain	54..61	
FT	Domain	/label= Loop_I	
FT	Domain	76..89	
FT	Domain	/label= Loop_II	
FT	Domain	92..100	
FT	Domain	/label= Loop_III	
FT	Domain	170..176	
FT	Domain	/label= Loop_IV	
FT	Domain	247..252	
FT	Domain	/label= Loop_V	
FT	Domain	266..277	
FT	Domain	/label= Loop_VI	
FT	Domain	280..288	
FT	Domain	/label= Loop_VII	
FT	Domain	362..368	
FT	Domain	/label= Loop_VIII	
FT	Domain	439..447	
FT	Domain	/label= Loop_IX	
FT	Domain	461..475	
FT	Domain	/label= Loop_X	
FT	Domain	478..486	
FT	Domain	/label= Loop_XI	
FT	Domain	560..566	

FT	/label= Loop_XII									
XX										
FN	WO200179443-A2.									
XX										
PD	25-OCT-2001.									
XX										
PF	12-APR-2001; 2001WO-US011924.									
XX										
PR	12-APR-2000; 2000US-0229358P.									
PR	25-APR-2000; 2000US-0199384P.									
PR	21-DEC-2000; 2000US-0256931P.									
XX										
PA	(HUMA-) HUMAN GENOME SCI INC.									
XX										
PI	Rosen CA, Haseltine WA;									
XX										
DR	WPI; 2001-616754/71.									
XX	N-PSDB; AAD21638.									
XX										
PT	Albumin fusion proteins comprising a therapeutic protein and albumin,									
PT	useful in the treating immune system disorders (e.g. transplant									
PT	rejection), blood related disorders (e.g. myocardial infarction) and									
PT	hyperproliferative disorders.									
XX										
PS	Claim 1; Fig 9; 380pp; English.									
XX										
CC	The invention relates to albumin fusion proteins comprising therapeutic									
CC	protein and human albumin (HA). Therapeutic proteins fused to albumin have									
CC	an extended shelf-life. The albumin fusion proteins are useful in the									
CC	treatment, prevention, diagnosis and/or detection of diseases, disorders									
CC	such as immune system disorders (e.g. transplant rejection), blood									
CC	related disorders (e.g. myocardial infarction), hyperproliferative									
CC	disorders (e.g. childhood acute myeloid leukaemia), renal disorders (e.g.									
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),									
CC	respiratory disorders (e.g. non-allergic rhinitis), neurological diseases									
CC	(e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),									
CC	reproductive system disorders (e.g. syphilis), infectious diseases (e.g.									
CC	measles), gastrointestinal disorders (e.g. irritable bowel syndrome) and									
CC	wound healing. Nucleic acids encoding albumin fusion protein is used in									
CC	gene therapy. The present sequence is human albumin (HA) protein									
XX										
SQ	Sequence 585 AA;									
	Query Match	100.0%;	Score 3103;	DB 4;	Length 585;					
	Best Local Similarity	100.0%;	Pred. No. 1.9e-254;							
	Matches 585;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1	DAHKSEVAHRFKDLGEENFKALVLI	FAQY	LQOC	PFEDHVKLVNEVTEFAKTCVADES	AE	60			
Db	1	DAHKSEVAHRFKDLGEENFKALVLI	FAQY	LQOC	PFEDHVKLVNEVTEFAKTCVADES	AE	60			
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNP	LPRLVRPEV	120						
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNP	LPRLVRPEV	120						
QY	121	DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP	180							
Db	121	DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP	180							
QY	181	KLDELRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEPAEVS	KLVTDLTK	240						
Db	181	KLDELRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEPAEVS	KLVTDLTK	240						
QY	241	VHTECHGDLLECCADDRADLAKYICENQDSSISKLKECEKPLLEKSHCIAEVENDEMPA	300							
Db	241	VHTECHGDLLECCADDRADLAKYICENQDSSISKLKECEKPLLEKSHCIAEVENDEMPA	300							
QY	301	DLPSLAADFVESKOVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRSLAKTYETTTLEKC	360							
Db	301	DLPSLAADFVESKOVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRSLAKTYETTTLEKC	360							
QY	361	CAAADPHCYAKVDFEFKPLVEEPQNLIKONCELFQQLGEYKFNQALLVRYTKKVPQVST	420							

Db 361 CAADPHECYAKVDFEFKPLVBPQNLIKQNCLEFQLGGEYKFNQALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 540  
QY 541 KEQLKAVMDDDFAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAFVEKCKCKADDKETCFAEEGKLVAAASQAALGL 585

RESULT 11  
AAE13399  
ID AAE13399 standard; protein; 585 AA.  
XX  
AC AAE13399;  
XX  
DT 12-FEB-2002 (first entry)  
XX  
DE Human albumin (HA) protein.  
XX  
KW Human; albumin; HA; fusion protein; immune system disorder; syphilis;  
KW transplant rejection; blood related disorder; myocardial infarction;  
KW hyperproliferative disorder; acute myeloid leukaemia; renal disorder;  
KW glomerulonephritis; cardiovascular disorder; arrhythmia; rhinitis;  
KW respiratory disorder; neurological disease; Alzheimer's disease;  
KW endocrine disorder; pheochromocytoma; reproductive system disorder;  
KW measles; gastrointestinal disorder; irritable bowel syndrome; HIV;  
KW human immunodeficiency virus; wound healing; renal cell carcinoma;  
KW melanoma; gene therapy.

OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 54..61  
FT /label= Loop\_I  
FT Domain 76..89  
FT /label= Loop\_II  
FT Domain 92..100  
FT /label= Loop\_III  
FT Domain 170..176  
FT /label= Loop\_IV  
FT Domain 247..252  
FT /label= Loop\_V  
FT Domain 266..277  
FT /label= Loop\_VI  
FT Domain 280..288  
FT /label= Loop\_VII  
FT Domain 362..368  
FT /label= Loop\_VIII  
FT Domain 439..447  
FT /label= Loop\_IX  
FT Domain 461..475  
FT /label= Loop\_X  
FT Domain 478..486  
FT /label= Loop\_XI  
FT Domain 560..566  
FT /label= Loop\_XII  
XX  
PN WO200179258-A1.  
PD 25-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US012008.  
XX  
PR 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
PR 21-DEC-2000; 2000US-0256931P.  
XX

(HUMA-) HUMAN GENOME SCI INC.  
(PRIN-) PRINCIPIA PHARM CORP.  
Rosen CA, Sadeghi H, Prior CP, Turner AJ;  
WPI: 2001-602931/68.  
N-PSDB; AAD22287.  
Albumin fusion proteins comprising a therapeutic protein and albumin,  
PT useful in the treating metastatic renal cell carcinoma, metastatic  
PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human  
PT immunodeficiency virus) or infection.  
XX  
PS Claim 1; Fig 9; 325pp; English.  
XX  
CC The invention relates to albumin fusion proteins comprising therapeutic  
CC protein and human albumin (HA). The albumin fusion proteins are useful in  
CC the treatment, prevention, diagnosis, and/or detection of diseases;  
CC disorders such as immune system disorders (transplant rejection); blood  
CC related disorders (myocardial infarction); hyperproliferative disorders  
CC (childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);  
CC cardiovascular disorders (arrhythmias); respiratory disorders (non-  
CC allergic rhinitis); neurological diseases (Alzheimer's disease);  
CC endocrine disorders (pheochromocytoma); reproductive system disorders  
CC (syphilis); infectious diseases (measles); gastrointestinal disorders  
CC (irritable bowel syndrome) and wound healing. The albumin fusion proteins  
CC are also used in the treatment of metastatic renal cell carcinoma,  
CC metastatic melanoma, malignant melanoma and HIV (human immunodeficiency  
CC virus) infection. Nucleic acid encoding albumin fusion protein is useful  
CC in gene therapy. The present sequence is human albumin (HA) protein  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPFEDHVKLNVNEVTEFAKTCVADESAE 60  
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPFEDHVKLNVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVRPEV 120  
QY 121 DVMCTAFHDNERTFLKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 121 DVMCTAFHDNERTFLKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
QY 181 KLDELRLDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
Db 181 KLDELRLDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCECEKELLESKSHCIAEVNDEMPA 300  
Db 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLEKCECEKELLESKSHCIAEVNDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLRLRAKYETTTLEK 360  
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLRLRAKYETTTLEK 360  
QY 361 CAADPHECYAKVDFEFKPLVBPQNLIKQNCLEFQLGGEYKFNQALLVRYTKVPQVST 420  
Db 361 CAADPHECYAKVDFEFKPLVBPQNLIKQNCLEFQLGGEYKFNQALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKGHPKAT 540

```

Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585

RESULT 12
ID ABB79006
XX ABB79006 standard; protein; 585 AA.
AC ABB79006;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human mature albumin protein SEQ ID NO:18.
XX
KW Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW albumin fusion protein; cytostatic; anorectic; immunosuppressive;
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;
KW type I diabetes mellitus; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..194
FT /label= 1
FT Domain 1..105
FT /label= subdomain
FT Disulfide-bond 53..62
FT Disulfide-bond 75..91
FT Disulfide-bond 90..101
FT Region 106..119
FT /note= "flexible inter-subdomain linker region"
FT Domain 120..194
FT /label= subdomain
FT Disulfide-bond 124..169
FT Disulfide-bond 168..177
FT Domain 195..367
FT /label= 2
FT Domain 195..291
FT /label= subdomain
FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region 292..315
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT /label= subdomain
FT Disulfide-bond 392..438
FT Disulfide-bond 437..448
FT Disulfide-bond 461..477
FT Disulfide-bond 476..487
FT Region 492..511
FT /note= "flexible inter-subdomain linker region"
FT Domain 512..585
FT /label= subdomain
FT Disulfide-bond 514..559
FT Disulfide-bond 558..567
XX
WO200179442-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011850.
XX
PR 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-0199384P.

```

```

PR 21-DEC-2000; 2000US-0256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2001-611723/70.
DR N-PSDB; ABN87288.
XX
PT New albumin fusion proteins, useful for treating diseases and disorders
PT such as cancer, comprise therapeutic protein fused to albumin.
XX
PS Claim 1; Fig 11; 413pp; English.
XX
CC The present invention describes an albumin fusion protein (I) comprising
CC a therapeutic protein: X and (a fragment or variant of) albumin
CC comprising a fully defined sequence in ABB79006 of 585 amino acids,
CC (where the fragment or variant has albumin or therapeutic protein: X
CC activity). (I) can have cytostatic, anorectic, immunosuppressive,
CC antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
CC Albumin fusion proteins are stabilised therapeutic proteins e.g.
CC antibodies to CS, C242 and CD80 useful for treating various diseases and
CC disorders such as non-Hodgkin's lymphoma, cancer, obesity, transplant
CC rejection, type I diabetes mellitus, rheumatoid arthritis and psoriasis.
CC Fusing albumin to therapeutic proteins stabilises the therapeutic
CC protein, extends the shelf life and retains the in vitro or in vivo
CC biological activity. It also reduces the need to formulate protein
CC solutions with large excesses of carrier proteins to prevent loss of
CC therapeutic proteins due to factors such as binding to the container. The
CC fusion proteins are easily dispensed with a simple formulation requiring
CC minimal post storage manipulation. The fusion of therapeutic proteins to
CC albumin confers stability in aqueous or other solution. The present
CC sequence represents the mature human albumin (HA) protein which is used
CC in the exemplification of the present invention
XX
SQ Sequence 585 AA;

```

```

Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.9e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNVEVTFKTCVADESAE 60
Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNVEVTFKTCVADESAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120

Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELFFAKRYKAAFTCCQAADKAACLLP 180

Qy 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 240

Qy 241 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVNDMPA 300
Db 241 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVNDMPA 300

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360

Qy 361 CAADPHECYAKVDFEPKPLVEEPQNLIKQNCELFEQLGEYKFQNALVRYTKVPQVST 420
Db 361 CAADPHECYAKVDFEPKPLVEEPQNLIKQNCELFEQLGEYKFQNALVRYTKVPQVST 420

Qy 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEDVLSVVLNQLCVLHKETPVSDRVTKCTES 480
Db 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEDVLSVVLNQLCVLHKETPVSDRVTKCTES 480

```



QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540  
DB LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585  
DB KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 13  
AAE08578  
ID AAE08578 standard; protein; 585 AA.  
XX AAE08578;  
XX  
DT 19-NOV-2001 (first entry)  
XX Human serum albumin (HSA).  
DE Human; albumin; cancer; cell proliferation; drug screening; biopsy.  
XX Homo sapiens.  
XX US6274305-B1.  
XX  
PD 14-AUG-2001.  
XX  
PF 19-DEC-1996; 96US-00769746.  
XX  
PR 19-DEC-1996; 96US-00769746.  
XX (TUFT ) UNIV TUFTS.  
XX  
XX Sonnenschein C. Soto AM;  
XX  
XX WPI: 2001-540371/60.  
DR N-PSDB; AAD11488.  
XX  
PT Measuring human cell proliferation, useful in drug screening to determine  
PT the potential for inhibiting cancer cell proliferation and for evaluating  
PT biopsied tumors, comprises employing albumin-derived peptide.  
XX  
XX Claim 1; Fig 1; 20pp; English.  
XX  
XX The invention related to a method for testing cancer cells. The method is  
CC useful for measuring human cancer cell proliferation, particularly for  
CC determining the potential for inhibiting cancer cells proliferation using  
CC albumin-derived peptides. The invention is also useful for drug screening  
CC assays, as well as for evaluating biopsied tumours. The present sequence  
CC is human serum albumin (HSA) related to the invention  
XX  
SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 4; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFDLGEENFKALVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 1 DAKSEVAHRFDLGEENFKALVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKLYEYARRHPYFAYAPPELLFFAKRYKAAFTCCQAADKACLIP 180  
DB 121 DVMCTAFHNEETFLKKLYEYARRHPYFAYAPPELLFFAKRYKAAFTCCQAADKACLIP 180  
QY 181 KLDELDEGKASSAKORLKACSLQKFGERAFKAWAVARLSQRPFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKACSLQKFGERAFKAWAVARLSQRPFAEVSKLVTDLTK 240

QY 241 VHTECCHGDLLECADDRAADLAKYICENQDSISSKLEKCECKPLEKSHCIAEVENDEMPA 300  
DB VHTECCHGDLLECADDRAADLAKYICENQDSISSKLEKCECKPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSTLAADFEVESKDVCKNYAEAKDVFGLMFLYAEARRHPDYSVVLRLAKTYETTTLEK 360  
DB DLPSTLAADFEVESKDVCKNYAEAKDVFGLMFLYAEARRHPDYSVVLRLAKTYETTTLEK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCELFEQLGGEYKFNQALLVRYTKKVPQVST 420  
DB CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCELFEQLGGEYKFNQALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDVLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB PTLVEVSRNLGVSKCKCHPEAKRMPCAEDVLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540  
DB LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585  
DB KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 14  
AAU75220  
ID AAU75220 standard; protein; 585 AA.  
XX AAU75220;  
XX  
DT 21-MAY-2002 (first entry)  
XX Mature form of human serum albumin (HSA or HA).  
XX  
XX Albumin fusion protein; therapeutic protein; immune disorder;  
KW autoimmune disorder; blood-related disorder; hyperproliferative disorder;  
KW renal disorder; cardiovascular disorder; respiratory disorder;  
KW neurological disorder; endocrine disorder; reproductive system disorder;  
KW gastrointestinal disorder; infectious disease; wound healing;  
KW human serum albumin; HSA; HA.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Protein 1..585  
FT /label= Mature\_HSA  
XX  
PN WO200179271-A1.  
XX  
PD 25-OCT-2001.  
XX  
XX 12-APR-2001; 2001WO-US012009.  
XX  
XX 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
PR 21-DEC-2000; 2000US-0256931P.  
XX  
XX (PRIN-) PRINCIPIA PHARM CORP.  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
XX Ballance DV, Sleep D, Turner AJ, Sadeghi H, Prior CP;  
XX WPI: 2002-179329/23.  
DR N-PSDB; ABK13862.  
XX  
XX New albumin fusion proteins with extended shelf life, useful for treating  
PT leukemia, warts, hepatitis, multiple sclerosis and AIDS, comprises  
PT therapeutic protein fused to albumin.  
XX  
XX Claim 1; Fig 15; 338pp; English.  
XX  
XX The present invention relates to albumin fusion proteins comprising a

CC therapeutic protein and albumin (e.g. human serum albumin (HSA or HA)).  
CC The albumin fusion proteins are useful for treating, preventing, or  
CC ameliorating various disorders. Such disorders include immune disorders,  
CC autoimmune disorders, blood-related disorders, hyperproliferative  
CC disorders, renal disorders, cardiovascular disorders, respiratory  
CC disorders, neurological disorders, endocrine disorders, reproductive  
CC system disorders. Gastrointestinal disorders, infectious disease, and  
CC wound healing. Therapeutic proteins can be stabilised to extend shelf  
CC life and/or retain the protein's activity for extended periods of time in  
CC solution, in vivo or in vitro by genetically or chemically fusing the  
CC protein to albumin or its fragment or variant. In addition the use of  
CC albumin fusion proteins reduces the need to formulate protein solutions  
CC with large excesses of carrier proteins to prevent loss of therapeutic  
CC protein due to factors such as binding to the container. The extension of  
CC shelf life was tested by measuring biological activity (Nb2 cell  
CC proliferation) of human albumin-human growth hormone (HA-hGH) fusion  
CC protein remaining after incubation in cell culture media for up to 3  
CC weeks. At week 3 there was still approximately 95% cell proliferation  
CC compared to no activity of unfused hGH. The present sequence represents  
CC the mature form of HSA which can be used to produce the albumin fusion  
CC proteins of the invention

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHLLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 120  
DB 61 NCDKSLHLLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 120  
QY 121 DVMCTAFHNDNETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
DB 121 DVMCTAFHNDNETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
QY 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVSKDKVCKNYAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLK 360  
QY 361 CAADPHECYAKVDFEKPVLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
DB 361 CAADPHECYAKVDFEKPVLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420  
QY 421 PTLVEVSNLGVSKGCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSNLGVSKGCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDYTPVKFENAEFTFFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
DB 481 LVNRRPCFSALEVDYTPVKFENAEFTFFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEBGKKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEBGKKLVAAASQAALGL 585

RESULT 15  
ABJ00986  
ID ABJ00986 standard; protein; 585 AA.  
XX

AC ABJ00986;  
XX  
DT 05-SEP-2002 (first entry)  
XX  
DE B lymphocyte stimulator protein binding peptide related protein.  
XX  
DE B lymphocyte stimulator protein binding protein; Blys; immune disease;  
KW allergy; proliferative disease; infectious disease; arteriosclerosis;  
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;  
KW ischaemia; graft-versus-host disease; neurodegenerative disease;  
KW immunosuppressive; nephrotropic; antirheumatic; antiarthritis;  
KW neuroprotective; cytosolic; immunostimulant; antitumour; anti-HIV;  
KW antiasthmatic; antiallergic; thymomimetic; antianemic; haemostatic;  
KW dermatological; antiinflammatory; cardiant; ophthalmological; uropathic;  
KW antidiabetic; antithyroid; antidepressant; hepatotropic.  
XX  
OS Homo sapiens.  
XX  
OS WO200216411-A2.  
PN  
XX 28-FEB-2002.  
PD  
XX  
XX 17-AUG-2001; 2001WO-US025850.  
PF  
XX 18-AUG-2000; 2000US-0226700P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Beltzer JP, Potter DM, Fleming TL, Rosen CA;  
PI  
XX WPI; 2002-499775/53.  
XX  
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises  
PT administering B lymphocyte stimulator binding polypeptide.  
PT  
XX  
PS Disclosure; Page 379-382; 387pp; English.  
XX  
XX The present invention relates to the treatment, prevention or  
CC amelioration of a disease or disorder associated with: aberrant B  
CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells  
CC of haematopoietic origin; or proliferative disease; and reducing,  
CC inhibiting or stimulating immunoglobulin production, B cell proliferation  
CC and graft rejection involving administration of Blys binding polypeptide.  
CC The Blys binding polypeptides are used in the treatment, prevention or  
CC amelioration of diseases such as immune system diseases, proliferative  
CC diseases, diseases of cells of hematopoietic origin, graft rejection,  
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,  
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and  
CC neurodegenerative diseases. The present sequence is a protein described  
CC in the invention  
XX  
SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHLLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 120  
DB 61 NCDKSLHLLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 120  
QY 121 DVMCTAFHNDNETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
DB 121 DVMCTAFHNDNETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
QY 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240



XX 28-FEB-2002.  
XX  
XX  
XX 17-AUG-2001; 2001WO-US025891.  
XX  
XX 18-AUG-2000; 2000US-0226489P.  
XX  
XX (DYAX-) DVAX CORP.  
XX  
XX Beltzer JP, Potter MD, Fleming TJ, Ladner RC;  
XX WPI; 2002-351647/38.  
XX  
XX New B-lymphocyte stimulator binding polypeptide useful in detecting or  
PT isolating BlyS or BlyS-like polypeptide comprises a specified amino acid  
PT sequence.  
XX  
XX Disclosure; Page 261-264; 269pp; English.  
XX  
XX The invention relates to a B lymphocyte stimulator (BlyS) binding  
CC polypeptide. BlyS binding peptides bind BlyS or BlyS-like proteins  
CC reversibly or irreversibly. The binding peptides are used in detection,  
CC isolation and/or purification of BlyS in a solution such as water or a  
CC buffer solution, as well as any fluid and/or cell obtained from an  
CC individual biological fluid, body tissue, body cell, cell line, tissue  
CC culture or other source containing BlyS or BlyS-like polypeptides. The  
CC biological fluids include sera, plasma, lymph, blood, blood fraction,  
CC urine, synovial fluid, spinal fluid, saliva and mucus. Sequences  
CC ABG33576, ABG33577 and ABG33847 represent human B lymphocyte stimulator  
CC proteins  
XX  
XX Sequence 585 AA;  
SQ

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180  
QY 181 KLDELURDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
DB 181 KLDELURDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRKAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRKAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEPLVEEPQNLKQNCLEFQGEYKFNALLVRYTKVPQVST 420  
DB 361 CAAADPHECYAKVDFEPLVEEPQNLKQNCLEFQGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCTSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
DB 481 LVNRRPCTSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540

QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 18  
ABG71291  
ID ABG71291 standard; protein; 585 AA.  
XX  
XX ABG71291;  
AC  
XX 08-JAN-2003 (first entry)  
DT  
XX Glycosylated protein determination associated protein.  
DE  
XX Standard substance; accuracy control substance; glycosylated protein;  
XX glycosylated albumin; fructosamine; diabetes; antidiabetic.  
KW  
XX Unidentified.  
OS  
XX JP2002243731-A.  
PN  
XX 28-AUG-2002.  
PD  
XX 21-FEB-2001; 2001JP-00045085.  
PF  
XX 21-FEB-2001; 2001JP-00045085.  
PR  
XX (KOKU-) KOKUSAI SHYAKU KK.  
XX (YOSH-) YOSHITOMI PHARM IND KK.  
PA  
XX WPI; 2002-744850/81.  
DR  
XX  
XX  
XX A standard substance for determination of glycosylated protein including  
PT glycosylated albumin and fructosamine, used in diagnosis of diabetes.  
PT  
XX Disclosure; Page 4; 6pp; Japanese.  
PS  
XX The present invention relates to a new standard and accuracy control  
CC substance for determination of glycosylated protein. The invention is  
CC useful for determination of glycosylated protein in the diagnosis of  
CC diabetes. Glycosylated albumin and fructosamine provide favourable  
CC dilution linearity. The present amino acid sequence represents the  
CC glycosylated protein determination associated protein as described in the  
CC invention  
XX  
XX Sequence 585 AA;  
SQ

Query Match 100.0%; Score 3103; DB 5; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLIP 180  
QY 181 KLDELURDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
DB 181 KLDELURDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRKAKTYETTLK 360

Db 301 DLPSLAADFVSKDVKCKYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
Qy 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540  
Qy 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585

RESULT 19  
ABR55695  
ID ABR55695 standard; protein; 585 AA.  
XX  
AC ABR55695;  
XX  
XX  
DT 18-AUG-2003 (first entry)  
XX  
DE Human albumin (HA) mature polypeptide.  
XX  
KW Albumin; HA; cytostatic; antibacterial; virucide; fungicide; anti-HIV;  
KW antialthmatic; osteopathic; antiarthritic; antiinflammatory; nootropic;  
KW neuroprotective; anti-thyroid; anti-ulcer; hepatotropic; vulnerary;  
XX  
OS Homo sapiens.  
XX  
XX WO2003030821-A2.  
PN  
PD 17-APR-2003.  
XX  
PF 04-OCT-2002; 2002WO-US031794.  
XX  
PR 05-OCT-2001; 2001US-0327281P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Haseeltine WA;  
PI  
XX WPI; 2003-468174/44.  
DR N-PSDB; ACC78574.  
XX  
XX New albumin fusion protein comprising a therapeutic protein:X, and  
PT albumin, its variant or fragment, useful for treating a cancer, AIDS  
PT asthma, leukemia, sepsis, endometriosis, osteoporosis, atherosclerosis,  
PT autism, or emphysema.  
XX  
PS Claim 1; Fig 15A-D; 455pp; English.  
XX  
CC The invention relates to an albumin fusion protein comprising a  
CC therapeutic protein:X, and albumin, its variant or fragment. The albumin  
CC fusion protein has the formula R1-L-R2; R2-L-R1; or R1-L-R2-L-R1 where R1  
CC is therapeutic protein:X or fragment, L is a peptide linker and R2 is  
CC albumin. The albumin fusion protein is useful for treating a disease or  
CC disorder that is modulated by therapeutic protein:X (claimed), such as  
CC cancer; infections (bacterial, viral, fungal, parasitic); or immune  
CC (AIDS, asthma); hematopoietic (leukemia, sepsis); reproductive (cystic  
CC fibrosis, endometriosis); musculoskeletal (osteoporosis, osteoarthritis);  
CC cardiovascular (congestive heart failure, atherosclerosis); neural/  
CC sensory (ataxia, attention deficit disorders, autism); respiratory  
CC (emphysema, bronchitis); endocrine (goiter, glomerulonephritis);  
CC digestive (ulcer, cirrhosis); or connective/epithelial (lupus, rheumatoid  
CC disorders. The present sequence represents a human albumin (HA) mature

CC polypeptide  
XX Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 6; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHRFKDLGEENFKALVLIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHSEVAHRFKDLGEENFKALVLIAFAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60  
Qy 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVRPEV 120  
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180  
Qy 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Db 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240  
Qy 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLUKCECKEPLLEKSHCIAEVEND 300  
Db 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLUKCECKEPLLEKSHCIAEVEND 300  
Qy 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
Db 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
Qy 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCFLFEQLGEYKFNALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKGKPKAT 540  
Qy 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585  
RESULT 20  
ABR42606  
ID ABR42606 standard; protein; 585 AA.  
XX  
AC ABR42606;  
XX  
DT 26-AUG-2003 (first entry)  
XX  
DE Human serum albumin.  
XX  
KW Human; serum albumin; abrogen; angiogenesis; inhibitor;  
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;  
KW gene therapy.  
OS Homo sapiens.  
XX  
PN WO2003042354-A2.  
PD  
XX 22-MAY-2003.  
XX  
PF 04-SEP-2002; 2002WO-US027885.  
XX  
PR 04-SEP-2001; 2001US-0316300P.  
XX

PA (AVET ) AVENTIS PHARM INC.  
XX  
PI Nesbit M, Fong TC, Brockstedt D;  
XX  
DR WPI; 2003-449566/42.  
XX  
PT New abrogen polypeptide, useful for treating an angiogenesis related  
XX diseases e.g. tumor metastasis.  
XX  
PS Disclosure; Page 26; 95pp; English.  
XX  
CC The present sequence is the protein sequence of human serum albumin  
CC (HSA). The invention provides novel abrogen polypeptides (see ABR42599-  
CC 602) that include a kringle domain, e.g. from urokinase plasminogen  
CC activator kringle domain. Abrogens are potent inhibitors of endothelial  
CC proliferation and angiogenesis, and have been shown to be capable of  
CC inhibiting or reducing cell proliferation induced by both basic  
CC fibroblast growth factor and vascular endothelial growth factor in a  
CC specific endothelial cell proliferation assay. The abrogen may be coupled  
CC to an N-terminal interleukin-2 signal peptide and to a C-terminal  
CC stabilising molecule such as HSA (see ABR42608-13). Abrogen polypeptides  
CC and polynucleotides are used in claimed methods of treating an  
CC angiogenesis-related disease or disorder, e.g. tumour metastasis  
XX  
SQ Sequence 585 AA;  
  
Query Match 100.0%; Score 3103; DB 7; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1.9e-254;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120  
  
QY 121 DVMCTAFHDNEETFLLKKYLIEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMCTAFHDNEETFLLKKYLIEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
  
QY 181 KLDELDRDGKASSAKQRLKCSLQKFGERAFKAMAVARLSQRFPAKAEVSKLVTDLTG 240  
DB 181 KLDELDRDGKASSAKQRLKCSLQKFGERAFKAMAVARLSQRFPAKAEVSKLVTDLTG 240  
  
QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADPVESKDYCKNYAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADPVESKDYCKNYAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360  
  
QY 361 CAADPHECYAKVDFEFPKPLVEEPONLIKQNCLEFEQLGEYKFQNALLVRYTKVPQVST 420  
DB 361 CAADPHECYAKVDFEFPKPLVEEPONLIKQNCLEFEQLGEYKFQNALLVRYTKVPQVST 420  
  
QY 421 PTLVEVSRNLGVSKCKCKHPKEMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVSKCKCKHPKEMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
  
QY 481 LVNRRPCFSALVEDTYPKBFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALVEDTYPKBFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
  
QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKLVLAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKLVLAASQAALGL 585







```

; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-702-572-2

Query Match 100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.5e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYLOQCPEDEHVHKLNVETFEAKTCVADESAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYLOQCPEDEHVHKLNVETFEAKTCVADESAE 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180
DB 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180
QY 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPILKSHCHCIAEVNDMPA 300
DB 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPILKSHCHCIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVKSGKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 421 PTLVEVSRNLGVKSGKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585

RESULT 3
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco

```

```

; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.5e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYLOQCPEDEHVHKLNVETFEAKTCVADESAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYLOQCPEDEHVHKLNVETFEAKTCVADESAE 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180
DB 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180
QY 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPILKSHCHCIAEVNDMPA 300
DB 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKLKECCCKPILKSHCHCIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVKSGKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 421 PTLVEVSRNLGVKSGKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKGHPKAT 540
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585

RESULT 4

```

```

US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/977
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 667
US-09-976-594-977

```

Query Match	100.0%	Score	3103;	DB	4;	Length	609;
Best Local Similarity	100.0%;	Pred.	No. 6.9e-287;				
Matches	585;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	DAHKSEVAHRRPKD	LGEEFNFKALVLI	IAFAQYLQCCPF	FEDHVKLVNEVTE	FAKTCVADSAE	60
Db							
Qy	25	DAHKSEVAHRRPKD	LGEEFNFKALVLI	IAFAQYLQCCPF	FEDHVKLVNEVTE	FAKTCVADSAE	84
Db							
Qy	61	NCDKSLHTLFGDKLCT	VATLTETTYGEMADCC	AKQEPERNECF	LQHKDDNP	LPRLVRREV	120
Db							
Qy	85	NCDKSLHTLFGDKLCT	VATLTETTYGEMADCC	AKQEPERNECF	LQHKDDNP	LPRLVRREV	144
Db							
Qy	121	DVMCTAHDNEETFL	KVYLGIARRHPHY	FYAPELLFFAKRY	KAAFTTECCQ	ADKAACLLP	180
Db							
Qy	145	DVMCTAHDNEETFL	KVYLGIARRHPHY	FYAPELLFFAKRY	KAAFTTECCQ	ADKAACLLP	204
Db							
Qy	181	KLDELDEGKASSAK	QRLKCLASLOKFG	ERAFKANAVARLS	QRFPKAAEFAEY	SKLVTDLTK	240
Db							
Qy	205	KLDELDEGKASSAK	QRLKCLASLOKFG	ERAFKANAVARLS	QRFPKAAEFAEY	SKLVTDLTK	264
Db							
Qy	241	VHTECCGHGDLLE	CADDRADLAKYI	CENODSIS	SKLKECCEKPL	LEKSHCIAE	300
Db							
Qy	265	VHTECCGHGDLLE	CADDRADLAKYI	CENODSIS	SKLKECCEKPL	LEKSHCIAE	324
Db							
Qy	301	DLPSLAADFVESK	DKVCKNYAAKDV	FLGMFLVEYARR	HPDYSVVL	LLFLAKTYETTL	360
Db							
Qy	325	DLPSLAADFVESK	DKVCKNYAAKDV	FLGMFLVEYARR	HPDYSVVL	LLFLAKTYETTL	384
Db							
Qy	361	CAAADPHECYAKV	DFEKPVEEPQNL	IKNCELFQ	ELGEYKFNQALL	VRVTKVPQVST	420
Db							
Qy	385	CAAADPHECYAKV	DFEKPVEEPQNL	IKNCELFQ	ELGEYKFNQALL	VRVTKVPQVST	444
Db							
Qy	421	PTLVVEVSRNLG	KVGSKCKCHPEAK	RMPCAEDY	LSVLNLQVLV	LHEKTPSVDRVTK	480
Db							
Qy	445	PTLVVEVSRNLG	KVGSKCKCHPEAK	RMPCAEDY	LSVLNLQVLV	LHEKTPSVDRVTK	504
Db							
Qy	481	LVNRRPCFSALE	VEDTYPVKEFN	FAETFTFHADI	CTLSEK	RQIKKOTALVELV	540
Db							
Qy	505	LVNRRPCFSALE	VEDTYPVKEFN	FAETFTFHADI	CTLSEK	RQIKKOTALVELV	564
Db							
Qy	541	KEQLKAVMDDFAA	FVEKCKCKAD	DKETCTFAE	BGKKLVAA	SAOALGL	585
Db							
Qy	565	KEQLKAVMDDFAA	FVEKCKCKAD	DKETCTFAE	BGKKLVAA	SAOALGL	609
Db							

RESULT 5  
US-09-919-039-370  
; Sequence 370, Application US/09919039  
; Patent No. 6727066  
; GENERAL INFORMATION:

```

; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370

```

Query Match	100.0.0%;	Score 3103;	DB 4;	Length 609;
Best Local Similarity	100.0.0%;	Pred. No. 6.9e-287;		
Matches 585;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	DAHKSEVAHFRKDLGEENFKALVLI	FAOAYLQOCPPEDHVKLNVETFAKTCVADES	60
Db	25	DAHKSEVAHFRKDLGEENFKALVLI	FAOAYLQOCPPEDHVKLNVETFAKTCVADES	84
Qy	61	NCDKSLHTLFGDKLCTVATLRETY	GYEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	120
Db	85	NCDKSLHTLFGDKLCTVATLRETY	GYEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	144
Qy	121	DVMCTAFHDNEETFLKKYLIEIARR	HPYFVAPPELLPFAKRYKAAFTCCOAAKAAACLLP	180
Db	145	DVMCTAFHDNEETFLKKYLIEIARR	HPYFVAPPELLPFAKRYKAAFTCCOAAKAAACLLP	204
Qy	181	KLDELDEGKASAKQRLKCSAQKQGER	AFKAWAVARLSORFPKAEFAEVS	KLVTDLTK 240
Db	205	KLDELDEGKASAKQRLKCSAQKQGER	AFKAWAVARLSORFPKAEFAEVS	KLVTDLTK 264
Qy	241	VHTECCCHGDLLECADRRADI	IAKYICENQDSISSKLEKCECKPPLLEKSHCIAEVENDEMPA	300
Db	265	VHTECCCHGDLLECADRRADI	IAKYICENQDSISSKLEKCECKPPLLEKSHCIAEVENDEMPA	324
Qy	301	DLPSLAADFVESKDVCNKYAEAKDVF	LGMFLYEYARRHPDYSVLLLR	LRAKTYETTLK 360
Db	325	DLPSLAADFVESKDVCNKYAEAKDVF	LGMFLYEYARRHPDYSVLLLR	LRAKTYETTLK 384
Qy	361	CAAADPHECYAKVFDEFPKLVIEEP	QNLIKQNCBELFEQLGKEYKQNAL	LLVRYTKKVPQVST 420
Db	385	CAAADPHECYAKVFDEFPKLVIEEP	QNLIKQNCBELFEQLGKEYKQNAL	LLVRYTKKVPQVST 444
Qy	421	PTLVEVSRNLGKVGSKCKKHPEAKR	MPCAEDYLSVVLNQLCVLHEKTPVSDRVT	KCCCTES 480
Db	445	PTLVEVSRNLGKVGSKCKKHPEAKR	MPCAEDYLSVVLNQLCVLHEKTPVSDRVT	KCCCTES 504
Qy	481	LWNRRPCFSALEVDETTVPKEFNAET	PTFHADICTLSEKERQIKKOTALVELVKHKPKAT	540
Db	505	LWNRRPCFSALEVDETTVPKEFNAET	PTFHADICTLSEKERQIKKOTALVELVKHKPKAT	564
Qy	541	KEOLKAYMDDFAAFVEKCKCKADDKET	CFABEGKKLVAASQAALGL	585
Db	565	KEOLKAYMDDFAAFVEKCKCKADDKET	CFABEGKKLVAASQAALGL	609

RESULT 6  
US-08-797-689-2  
; Sequence 2, Application US/08797689  
; Patent No. 5876969  
; GENERAL INFORMATION:  
; APPLICANT: Floor; Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guitton, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice

;; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
;; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
;; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
;; NUMBER OF SEQUENCES: 36  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSES: Rhone-Poulenc Rorer Inc.  
;; STREET: 500 Arcola Road, 3C43  
;; CITY: Collegeville  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19426  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: Macintosh  
;; OPERATING SYSTEM: System 7.1  
;; SOFTWARE: Word 5.1 (PatentIn)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/797,689  
;; FILING DATE: 31-JAN-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/256,927  
;; FILING DATE: 28-JUL-1994  
;; APPLICATION NUMBER: FR 92/01064  
;; FILING DATE: 31-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/FR93/00085  
;; FILING DATE: 28-JAN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith Ph.D., Julie K.  
;; REGISTRATION NUMBER: P-38,619  
;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 610 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-797-689-2  
Query Match 100.0%; Score 3103; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAOYLQOCPEDHVKLVNEVTEFAKTCVADESAB 60  
DB 25 DAHKEVAHRFKDLGEENFKALVLIAPAOYLQOCPEDHVKLVNEVTEFAKTCVADESAB 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPRLVLRPEV 144  
QY 121 DVMCTAFHNDNEFTLKKLYEYIARRHPYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
DB 145 DVMCTAFHNDNEFTLKKLYEYIARRHPYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204  
QY 181 KLDELURDEGKASAKORLKCASIQKGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
DB 205 KLDELURDEGKASAKORLKCASIQKGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECKEPLEKSHCIAEVENDEMPA 300  
DB 265 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECKEPLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADPVEKQVCKNVAEKQVPLGMFLYEARHPDYSVLLLRLLAKTYETTLK 360  
DB 325 DLPSLAADPVEKQVCKNVAEKQVPLGMFLYEARHPDYSVLLLRLLAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQEGYKFQNALLVRYTKKVPQVST 420

DB 305 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQEGYKFQNALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
DB 505 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 564  
QY 541 KEOLKAVMDDFAAFVEKCKADDDKTCFAEEGKKLVAAASQAALGL 585  
DB 565 KEOLKAVMDDFAAFVEKCKADDDKTCFAEEGKKLVAAASQAALGL 609  
RESULT 7  
US-09-984-186-2  
; Sequence 2, Application US/09984186  
; Patent No. 6866179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guitton, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,186  
; FILING DATE: 29-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2  
Query Match 100.0%; Score 3103; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-2

Query Match 100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAB 60
Db 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAB 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERCEFLQHKDDNPRLVLRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERCEFLQHKDDNPRLVLRPEV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCOAADKAACLIP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCOAADKAACLIP 204
QY 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 240
Db 205 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 264
QY 241 VHTCCHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK 384
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFEQLGEYKFQNALVRYTKVPQVST 420
Db 385 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFEQLGEYKFQNALVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKCHPEAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTTCCTES 480
Db 445 PTLVEVSRLGKVGSKCKCHPEAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTTCCTES 504
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEBGKGLVAASQAALGL 585
Db 565 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEBGKGLVAASQAALGL 609

RESULT 10
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Robanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAB 60
Db 203 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAB 262
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERCEFLQHKDDNPRLVLRPEV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERCEFLQHKDDNPRLVLRPEV 322
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCOAADKAACLIP 180
Db 323 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCCOAADKAACLIP 382
QY 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 240
Db 383 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTG 442
QY 241 VHTCCHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPPLLEKSHCIAEVENDEMPA 300
Db 443 VHTCCHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPPLLEKSHCIAEVENDEMPA 502
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK 360
Db 503 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK 562
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFEQLGEYKFQNALVRYTKVPQVST 420
Db 563 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFEQLGEYKFQNALVRYTKVPQVST 622
QY 421 PTLVEVSRLGKVGSKCKCHPEAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTTCCTES 480
Db 623 PTLVEVSRLGKVGSKCKCHPEAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTTCCTES 682
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKPKAT 540
Db 683 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKPKAT 742
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEBGKGLVAASQAALGL 585
Db 743 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEBGKGLVAASQAALGL 787

RESULT 11
US-08-797-689-16

```



```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16

Query Match      100.0%; Score 3103; DB 4; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQOCPEFHVHKLNVNEVTEFAKTCVADESAAE 60
Db 203 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQOCPEFHVHKLNVNEVTEFAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 322

QY 121 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 323 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 382

QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 383 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 442

QY 241 VITECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 443 VITECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 502

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTLK 360
Db 503 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTLK 562

QY 361 CAAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420
Db 563 CAAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 622

QY 421 PTLVEVSRNLGVKSGKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 623 PTLVEVSRNLGVKSGKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 682

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 540
Db 683 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 742

QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
Db 743 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 787

RESULT 13
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
;

;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match      99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQOCPEFHVHKLNVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDGLGEENFKALVLIAPAYLQOCPEFHVHKLNVNEVTEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPVP 144

QY 121 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 145 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204

QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264

QY 241 VITECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 265 VITECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTLK 384

QY 361 CAAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 420
Db 385 CAAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALVRYTKVPQVST 444

QY 421 PTLVEVSRNLGVKSGKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGVKSGKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 564

QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 609

RESULT 14
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Briereley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

```

ADDRESS: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9108Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred No. 1.7e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDELRLDEGKVSSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVSKDKVKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVSKDKVKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALIVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALIVRYTKVPQVST 444
QY 421 PTLVEVSRNLGVSKCKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVSKCKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVTPVKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540
DB 505 LVNRRPCFSALEVDVTPVKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 564
QY 541 KEQLKAVNDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585

565 KEQLKAVNDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609
RESULT 15
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 99.9%; Score 3099; DB 4; Length 609;
Best Local Similarity 99.8%; Pred No. 1.7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDELRLDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVSKDKVKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVSKDKVKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALIVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALIVRYTKVPQVST 444
QY 421 PTLVEVSRNLGVSKCKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVSKCKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVTPVKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540
DB 505 LVNRRPCFSALEVDVTPVKPEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 564
QY 541 KEQLKAVNDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
565 KEQLKAVNDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609
RESULT 16
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075



GENERAL INFORMATION:  
APPLICANT: AMGEN INC.  
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
NUMBER OF INVENTIONS: 33  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Center, Patent Operations/RRC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: U.S.  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04075  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;  
Best Local Similarity 99.8%; Pred. No. 1.7e-286;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144

QY 121 DVMCTAFHDNEBTFLLKLYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
DB 145 DVMCTAFHDNEBTFLLKLYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 204

QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLTK 264

QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCHIAEVNDMPA 300  
DB 265 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCHIAEVNDMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 384

QY 361 CAAADPHCYAKVDFEPLVEEPQNLIKONCELPQGEYKFQNALVRYTKVPQVST 420  
DB 385 CAAADPHCYAKVDFEPLVEEPQNLIKONCELPQGEYKFQNALVRYTKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
DB 505 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 564

QY 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAASQAALGL 585  
DB 565 KEOLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAASQAALGL 609

RESULT 17  
US-08-897-956A-3  
Sequence 3, Application US/08897956A  
Patent No. 64231512  
GENERAL INFORMATION:  
APPLICANT: Mary Ellen Digan  
APPLICANT: Philip Lake  
APPLICANT: Hermann Gram  
TITLE OF INVENTION: Fusion Polypeptides  
FILE REFERENCE: 600-7244/CPA  
CURRENT APPLICATION NUMBER: US/08/897,956A  
CURRENT FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/022,689  
PRIOR FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 978  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Fusion polypeptide  
US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;  
Best Local Similarity 99.8%; Pred. No. 8e-286;  
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 212 DAHKEVAHRFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 271

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120  
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 331

QY 121 DVMCTAFHDNEBTFLLKLYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
DB 332 DVMCTAFHDNEBTFLLKLYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 391

QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLTK 240  
DB 392 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPKAEFAEVSKLVTDLTK 451

QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCHIAEVNDMPA 300  
DB 452 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCHIAEVNDMPA 511

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360  
DB 512 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 571

QY 361 CAAADPHCYAKVDFEPLVEEPQNLIKONCELPQGEYKFQNALVRYTKVPQVST 420  
DB 572 CAAADPHCYAKVDFEPLVEEPQNLIKONCELPQGEYKFQNALVRYTKVPQVST 631

QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCCTES 480  
DB 632 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVVNLQCVLHEKTPVSDRVTKCCTES 691

QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 540  
DB 692 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKGHPKAT 751

QY 541 KEOLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAASQAALGL 584  
DB 752 KEOLKAVMDDDFAAFVEKCKCKADDKETCFABEKGKLVAAASQAALGL 795

RESULT 18  
US-08-448-196A-3

; Sequence 3, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;  
Best Local Similarity 99.7%; Pred. No. 5.8e-286;  
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
  
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
  
QY 181 KLDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240  
Db 181 KLDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240  
  
QY 241 VHTECCGHDLLDECADRDADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
Db 241 VHTECCGHDLLDECADRDADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
  
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCPELQGEYKFQNALLVRYTKKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCPELQGEYKFQNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVWLNQLCVLHKEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVWLNQLCVLHKEKTPVSDRVTKCCTES 480  
  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
Db 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
  
QY 541 KEQLKAVMDDPFAAFAVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDPFAAFAVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
  
RESULT 19  
US-08-984-176-1  
; Sequence 1, Application US/08984176  
; Patent No. 5948609  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C  
; APPLICANT: HO, JOSEPH X  
; APPLICANT: RUKER, FLORIAN  
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER  
; FILE REFERENCE: 08/984,176  
; CURRENT APPLICATION NUMBER: US/08/984,176  
; CURRENT FILING DATE: 1997-12-03  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;  
Best Local Similarity 99.7%; Pred. No. 5.8e-286;  
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120  
  
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
  
QY 181 KLDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240  
Db 181 KLDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240  
  
QY 241 VHTECCGHDLLDECADRDADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
Db 241 VHTECCGHDLLDECADRDADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTTLEK 360  
  
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCPELQGEYKFQNALLVRYTKKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCPELQGEYKFQNALLVRYTKKVPQVST 420  
  
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVWLNQLCVLHKEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVWLNQLCVLHKEKTPVSDRVTKCCTES 480  
  
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVKKPKAT 540  
Db 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQTALVELVKKPKAT 540

Db 481 LVNRRPCFSALEVDYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540

QY 541 KEOLKAVMDDFAAAFVCKCKKADDDKTCFAEEGKKLVAASQAALGL 585

Db 541 KEOLKAVMDDFAAAFVCKCKKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 20

US-08-448-196A-5

Sequence 5, Application US/08448196A

Patent No. 5780594

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.

TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

TITLE OF INVENTION: RELATED PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: NASA

STREET: MARSHALL SPACE FLIGHT CENTER

CITY: HUNTSVILLE

STATE: ALABAMA

COUNTRY: USA

ZIP: 35812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/448,196A

FILING DATE: 23-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: BROAD JR., ROBERT L.

REGISTRATION NUMBER: 18,757

REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 205-544-0021

TELEFAX: 205-544-0258

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 583 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

US-08-448-196A-5

Query Match 79.2%; Score 2458.5; DB 1; Length 583;

Best Local Similarity 75.8%; Pred. No. 1.7e-225;

Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHKSEVAHRFKDLGEENFKALVIAFAOYLQCCPEDHVKLVNEVTEPAKTCVADESAB 60

Db 1 DTHKSEIAHRFDLGEKFKGLVAFSQYLQCCPEDHVKLVNEVTEPAKCAADESAB 60

QY 61 NCCKSLHTLFGDKLCTVATRLTYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV 120

Db 61 NCCKSLHTLFGDKLCTVATRLTYGELADCCQEPERNECFLTHKDHNPRLK-KPEP 119

QY 121 DVMCTAFHDSNETFLKKLYLAEARRHPFYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

Db 120 DAQCAAFQEDPKFLGKLYLEVARRHPFYFGPELLFHAEEYKADFTCCPADDKLCLIP 179

QY 181 KLDELURDEGKASAKORLKASLQKFGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

Db 180 KLDALKERILLSSAKERLKCSSQNFGERAVKAWVARLSQKFPKADFAEVSCLVTDLT 239

QY 241 VHTTECHGDLLECADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

Db 240 VHKCECHGDLLECADRADLAKYICEHQDSISGKLKACCDKPLLQKSHCIAEVEKEDDLS 299

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLYEARRHDPDYSVLLLRLAKTYYETTTLEK 360

Db 300 DIPALAADFAEDKEICKHYKDAKDVFLGTFLVEYSRRHDPDYSVLLLRLAKTYYETTTLEK 359

QY 361 CAAADPHCEYAKVDFEFKPLVEEPQNLKQNCLEPEQLGEYKFQNALLVRYTKKVPQVST 420

Db 360 CAEADPPACYRTVFDQFTPLVBEPSLVKKNCDDLFEVGEYDFQNALIVRYTKKAPQVST 419

QY 421 PTLVEVSRLGKVGSKCKKHPEAKRMPCAEDYLSVVNLNOLCVLHHEKTPVSDRVTKCCTES 480

Db 420 PTLVEIGRTLGVGSRCCCKLPESERLPCSENHLALANLVCVLHHEKTPVSEKTKCCTDS 479

QY 481 LVNRRPCFSALEVDYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540

Db 480 LAERRPCFSALELDDEGYPVKEFKAETFTFHADICTLPEDEKQIKKQSALAEVLVKHKPKAT 539

QY 541 KEOLKAVMDDFAAAFVCKCKKADDDKTCFAEEGKKLVAASQAAL 583

Db 540 KEOLKAVMDDFAAAFVCKCKKADDDKTCFAEEGKKLVAASQAAL 582

Search completed: October 13, 2005, 08:50:56

Job time : 27 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:43:42 ; Search time 169 Seconds  
(without alignments)  
1442.485 Million cell updates/sec

Title: US-10-816-042-18  
Perfect score: 3103  
Sequence: 1 DHAKEVAHRFDLGEENFK.....TCFAEEGKKLVAASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	9	US-09-929-552-2
2	3103	100.0	585	10	US-09-932-613-445
3	3103	100.0	585	10	US-09-984-010-26
4	3103	100.0	585	10	US-09-833-041-18
5	3103	100.0	585	10	US-09-833-117-18
6	3103	100.0	585	10	US-09-932-322-445
7	3103	100.0	585	10	US-09-832-501-18
8	3103	100.0	585	10	US-09-833-118-18
9	3103	100.0	585	11	US-09-833-245-18
10	3103	100.0	585	11	US-09-832-929-18
11	3103	100.0	585	14	US-10-153-604A-5

12	3103	100.0	585	14	US-10-319-263-1	Sequence 1, Appli
13	3103	100.0	585	14	US-10-319-263-2	Sequence 2, Appli
14	3103	100.0	585	14	US-10-414-469-1	Sequence 1, Appli
15	3103	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
16	3103	100.0	585	14	US-10-413-831-1	Sequence 1, Appli
17	3103	100.0	585	14	US-10-413-831-2	Sequence 2, Appli
18	3103	100.0	585	15	US-10-413-832-1	Sequence 1, Appli
19	3103	100.0	585	15	US-10-413-832-2	Sequence 2, Appli
20	3103	100.0	585	15	US-10-414-386-1	Sequence 1, Appli
21	3103	100.0	585	15	US-10-414-386-2	Sequence 2, Appli
22	3103	100.0	585	15	US-10-233-675A-11	Sequence 11, Appli
23	3103	100.0	585	15	US-10-462-262-26	Sequence 26, Appli
24	3103	100.0	585	15	US-10-425-000-31	Sequence 31, Appli
25	3103	100.0	585	15	US-10-424-999-11	Sequence 11, Appli
26	3103	100.0	585	15	US-10-602-141-3	Sequence 3, Appli
27	3103	100.0	585	17	US-10-816-042-18	Sequence 18, Appli
28	3103	100.0	585	17	US-10-922-142-18	Sequence 18, Appli
29	3103	100.0	585	17	US-10-775-180-327	Sequence 327, Appli
30	3103	100.0	585	17	US-10-932-104-18	Sequence 18, Appli
31	3103	100.0	585	18	US-10-775-204-1038	Sequence 1038, Ap
32	3103	100.0	585	20	US-11-033-766-1	Sequence 1, Appli
33	3103	100.0	585	20	US-11-033-766-2	Sequence 2, Appli
34	3103	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
35	3103	100.0	609	10	US-09-919-039-370	Sequence 370, Appli
36	3103	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
37	3103	100.0	609	14	US-10-365-623-23	Sequence 23, Appli
38	3103	100.0	609	17	US-10-775-180-379	Sequence 379, Appli
39	3103	100.0	609	18	US-10-775-204-1094	Sequence 1094, Ap
40	3103	100.0	610	9	US-09-984-186-2	Sequence 2, Appli
41	3103	100.0	610	14	US-10-237-667-2	Sequence 2, Appli
42	3103	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
43	3103	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
44	3103	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
45	3103	100.0	610	14	US-10-237-624-2	Sequence 2, Appli
46	3103	100.0	610	15	US-10-702-536-2	Sequence 2, Appli
47	3103	100.0	610	15	US-10-702-636-2	Sequence 2, Appli
48	3103	100.0	616	15	US-10-433-108-13	Sequence 13, Appli
49	3103	100.0	619	18	US-10-775-204-269	Sequence 269, App
50	3103	100.0	619	18	US-10-775-204-270	Sequence 270, App
51	3103	100.0	619	18	US-10-775-204-1337	Sequence 1337, Ap
52	3103	100.0	621	18	US-10-775-204-1334	Sequence 1334, Ap
53	3103	100.0	623	17	US-10-775-180-573	Sequence 573, App
54	3103	100.0	623	17	US-10-775-180-577	Sequence 577, App
55	3103	100.0	623	17	US-10-775-180-588	Sequence 588, App
56	3103	100.0	623	18	US-10-775-204-1558	Sequence 1558, Ap
57	3103	100.0	623	18	US-10-775-204-1563	Sequence 1563, Ap
58	3103	100.0	623	18	US-10-775-204-1589	Sequence 1589, Ap
59	3103	100.0	624	15	US-10-433-108-16	Sequence 16, Appli
60	3103	100.0	631	15	US-10-433-108-14	Sequence 14, Appli
61	3103	100.0	634	18	US-10-775-204-1536	Sequence 1536, Ap
62	3103	100.0	634	18	US-10-775-204-1538	Sequence 1538, Ap
63	3103	100.0	635	17	US-10-775-180-91	Sequence 91, Appli
64	3103	100.0	635	17	US-10-775-180-92	Sequence 92, Appli
65	3103	100.0	635	17	US-10-775-180-93	Sequence 93, Appli
66	3103	100.0	635	17	US-10-775-180-94	Sequence 94, Appli
67	3103	100.0	635	18	US-10-775-204-301	Sequence 301, App
68	3103	100.0	635	18	US-10-775-204-307	Sequence 307, App
69	3103	100.0	635	18	US-10-775-204-308	Sequence 308, App
70	3103	100.0	635	18	US-10-775-204-309	Sequence 309, App
71	3103	100.0	635	18	US-10-775-204-1644	Sequence 1644, Ap
72	3103	100.0	636	18	US-10-775-204-1537	Sequence 1537, Ap
73	3103	100.0	636	18	US-10-775-204-1539	Sequence 1539, Ap
74	3103	100.0	636	18	US-10-775-204-1560	Sequence 1560, Ap
75	3103	100.0	636	18	US-10-775-204-1594	Sequence 1594, Ap
76	3103	100.0	637	17	US-10-775-180-578	Sequence 578, App
77	3103	100.0	637	17	US-10-775-180-579	Sequence 579, App
78	3103	100.0	637	18	US-10-775-204-1564	Sequence 1564, Ap
79	3103	100.0	638	16	US-10-361-997-69	Sequence 1565, Ap
80	3103	100.0	638	18	US-10-775-204-1290	Sequence 1290, Ap
81	3103	100.0	638	18	US-10-775-204-1290	Sequence 1290, Ap
82	3103	100.0	638	18	US-10-503-834-69	Sequence 69, Appli
83	3103	100.0	639	17	US-10-775-180-129	Sequence 129, App
84	3103	100.0	639	17	US-10-775-180-130	Sequence 130, App

85 3103 100.0 639 17 US-10-775-180-131 Sequence 131, App  
86 3103 100.0 639 17 US-10-775-180-132 Sequence 132, App  
87 3103 100.0 639 18 US-10-775-204-414 Sequence 414, App  
88 3103 100.0 639 18 US-10-775-204-416 Sequence 416, App  
89 3103 100.0 639 18 US-10-775-204-417 Sequence 417, App  
90 3103 100.0 639 18 US-10-775-204-418 Sequence 418, App  
91 3103 100.0 639 18 US-10-775-204-1285 Sequence 1285, App  
92 3103 100.0 639 18 US-10-775-204-1295 Sequence 1295, App  
93 3103 100.0 640 15 US-10-433-108-15 Sequence 15, App1  
94 3103 100.0 640 15 US-10-433-108-17 Sequence 17, App1  
95 3103 100.0 640 17 US-10-775-180-615 Sequence 615, App  
96 3103 100.0 640 18 US-10-775-204-280 Sequence 280, App  
97 3103 100.0 640 18 US-10-775-204-288 Sequence 288, App  
98 3103 100.0 640 18 US-10-775-204-1602 Sequence 1602, App  
99 3103 100.0 640 18 US-10-775-204-1603 Sequence 1603, App  
100 3103 100.0 640 18 US-10-775-204-1627 Sequence 1627, App

ALIGNMENTS

RESULT 1  
US-09-929-552-2  
; Sequence 2, Application US/09929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Sonenschein, Carlos  
; Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/929,552  
; FILING DATE: 14-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/769,746  
; FILING DATE: 19-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-929-552-2  
Query Match 100.0%; Score 3103; DB 9; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60

QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120  
DB 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180  
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180  
QY 181 KLDELRLDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPPKAEFAEYVKLVTDLT 240  
DB 181 KLDELRLDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPPKAEFAEYVKLVTDLT 240  
QY 241 VHTECHGDLLECCADDRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTECHGDLLECCADDRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DPLSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLLLRLLAKTYETTLEK 360  
DB 301 DPLSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLLLRLLAKTYETTLEK 360  
QY 361 CAAADPHCYAKVDFDEKPLVEEPONLIKQNCBFLFEQLGEYKFNQALLVRYTKKVPQVST 420  
DB 361 CAAADPHCYAKVDFDEKPLVEEPONLIKQNCBFLFEQLGEYKFNQALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDVETVYVPEFNAETFTFHADICTLSEKEROIKKQTLALVELVHKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVYVPEFNAETFTFHADICTLSEKEROIKKQTLALVELVHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 2

US-09-932-613-445  
; Sequence 445, Application US/09932613  
; Publication No. US20030091565A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; APPLICANT: Beltzer, James P.  
; APPLICANT: Potter, M. Daniel  
; APPLICANT: Fleming, Tony J.  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON  
; FILE REFERENCE: Dyx-025.1 PCT; DYX-025.1 US  
; CURRENT APPLICATION NUMBER: US/09/932,613  
; CURRENT FILING DATE: 2001-08-17  
; NUMBER OF SEQ ID NOS: 458  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 445  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: HomoSapiens  
US-09-932-613-445

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120  
DB 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAAFTECCQAADKAAACLLP 180

Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKCAASLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240  
Db 181 KLDELDEGKASSAKQRLKCAASLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240  
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHGICIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHGICIAEVENDEMPA 300  
QY 301 DLPSSLAADPVESKDVCKNVAEAKDVLGMLFYEARRRHPDYSVLLRLRAKTYETTTLEKC 360  
Db 301 DLPSSLAADPVESKDVCKNVAEAKDVLGMLFYEARRRHPDYSVLLRLRAKTYETTTLEKC 360  
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEAGKGLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEAGKGLVAASQAALGL 585

RESULT 3

US-09-984-010-26  
; Sequence 26, Application US/09984010  
; Publication No. US20030104578A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David James  
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE  
; AND SERUM ALBUMIN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT & DUNNER, LLP  
; STREET: 1300 I Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,010  
; FILING DATE: 21-May-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/091,873  
; FILING DATE: 25-JUN-1998  
; APPLICATION NUMBER: PCT/GB96/03164  
; FILING DATE: 19-DEC-1996  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-984-010-26

Query Match 100.0%; Score 3103; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVVAHRFPKDLGLENFKAIVLAFQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHKEVVAHRFPKDLGLENFKAIVLAFQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDPNLPRLVRPEV 120  
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDPNLPRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKCAASLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240  
Db 181 KLDELDEGKASSAKQRLKCAASLOKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLVTDLT 240  
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHGICIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHGICIAEVENDEMPA 300  
QY 301 DLPSSLAADPVESKDVCKNVAEAKDVLGMLFYEARRRHPDYSVLLRLRAKTYETTTLEKC 360  
Db 301 DLPSSLAADPVESKDVCKNVAEAKDVLGMLFYEARRRHPDYSVLLRLRAKTYETTTLEKC 360  
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420  
Db 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
Db 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEAGKGLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEAGKGLVAASQAALGL 585

RESULT 4

US-09-833-041-18  
; Sequence 18, Application US/09833041  
; Publication No. US20030125247A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF545  
; CURRENT APPLICATION NUMBER: US/09/833,041  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-041-18

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DAHSEVAHRFKDOLGBENFKALVLJIAFAQYIQQCPFEDHVKLVNEVTFEAKTQVADESAB	60
Db	1	DAHSEVAHRFKDOLGBENFKALVLJIAFAQYIQQCPFEDHVKLVNEVTFEAKTQVADESAB	60
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQEPBERNECFLOHKDDNPRLVRPVE	120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQEPBERNECFLOHKDDNPRLVRPVE	120
Qy	121	DMVCTAFHDNEETFLLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP	180
Db	121	DMVCTAFHDNEETFLLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP	180
Qy	181	KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSORPPKAFPAEVSKLVTDLTK	240
Db	181	KLDELDEGKASAKQRLKCSLQKFGGERAFKAWAVARLSORPPKAFPAEVSKLVTDLTK	240
Qy	241	VHTECCHGDLLECCADRDADLAKYICENODSISSKLKCECKPALLEKSHCIAEVENDEMPA	300
Db	241	VHTECCHGDLLECCADRDADLAKYICENODSISSKLKCECKPALLEKSHCIAEVENDEMPA	300
Qy	301	DLPSLAADFVESKDVCKNYABAKOVFLGMFLFYEYARRHPDYSVVLLRLAKTYETTLKRC	360
Db	301	DLPSLAADFVESKDVCKNYABAKOVFLGMFLFYEYARRHPDYSVVLLRLAKTYETTLKRC	360
Qy	361	CAAADPHECYAKVDFEPLVEEPPONLIKQNCHELFEOLGEYKQONALLVRYTKKVPQVST	420
Db	361	CAAADPHECYAKVDFEPLVEEPPONLIKQNCHELFEOLGEYKQONALLVRYTKKVPQVST	420
Qy	421	PTLVEVSRNLGVGSKCKKHPEAKMPCABDYLSVNLQCLVHKEKTPVSDRVTCKCTES	480
Db	421	PTLVEVSRNLGVGSKCKKHPEAKMPCABDYLSVNLQCLVHKEKTPVSDRVTCKCTES	480
Qy	481	LVNRRPCFSALEVDETYVPKFENAEFTTFTHADICTLSEKERQIKKQATALVELVKHKPKAT	540
Db	481	LVNRRPCFSALEVDETYVPKFENAEFTTFTHADICTLSEKERQIKKQATALVELVKHKPKAT	540
Qy	541	KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL	585
Db	541	KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL	585

## RESULT 5

```

US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-833-117-18

```

Db	1	DABKSEVAHFPKDLGEENFKA	VLIIAFAQYLQCCPFEDHVKLVNEVTFB	AKTCVADESAB	60
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLQHKDDNPMLPRLVRREV	120		
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLQHKDDNPMLPRLVRREV	120		
Qy	121	DVMCTAFPHDNEETFLKKYLIEIARRHPYFVAPPELLFFAKRYKAAFTTECCQAADKAACLLP	180		
Db	121	DVMCTAFPHDNEETFLKKYLIEIARRHPYFVAPPELLFFAKRYKAAFTTECCQAADKAACLLP	180		
Qy	181	KLDELDRDEGKASSAKORLKCASLOKQGERAFKAWAVARLSORFPKABEFAEVS	KLVTDLTK	240	
Db	181	KLDELDRDEGKASSAKORLKCASLOKQGERAFKAWAVARLSORFPKABEFAEVS	KLVTDLTK	240	
Qy	241	VHTECCHGDLLECCADDRADIAKYICENQDSTISSKLECCCKPLLEKSHCIAEVENDEMPA	300		
Db	241	VHTECCHGDLLECCADDRADIAKYICENQDSTISSKLECCCKPLLEKSHCIAEVENDEMPA	300		
Qy	301	DLPSLAADFVESKDVCKNYAEAKDVFLGMPFIYEYARRHPDYSVLLLLRAKTYETTLK	360		
Db	301	DLPSLAADFVESKDVCKNYAEAKDVFLGMPFIYEYARRHPDYSVLLLLRAKTYETTLK	360		
Qy	361	CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCSELFQELGEYKFFQNALLVRVTKKVPQVST	420		
Db	361	CAAADPHECYAKVFDEFKPLVEEPQNLIKQNCSELFQELGEYKFFQNALLVRVTKKVPQVST	420		
Qy	421	PTLVEVSRNLGKVGSKCKKHEAKRMPCAEDYLSVLLNQLCVLHKEKTPVSDRVT	KCCCTES	480	
Db	421	PTLVEVSRNLGKVGSKCKKHEAKRMPCAEDYLSVLLNQLCVLHKEKTPVSDRVT	KCCCTES	480	
Qy	481	LVNRRPCFSALVDEVTVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKHPKAT	540		
Db	481	LVNRRPCFSALVDEVTVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKHPKAT	540		
Qy	541	KEOLKAYMDDFAAFVEKCCCKADDKETCFABEGKKLVAASQAALGL	585		
Db	541	KEOLKAYMDDFAAFVEKCCCKADDKETCFABEGKKLVAASQAALGL	585		

## RESULT 6

```

US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
; US-09-932-322-445

```



QY 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECCEKPLLEKSHCHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECCEKPLLEKSHCHCIAEVENDEMPA 300  
QY 301 DLPSSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360  
DB 301 DLPSSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360  
QY 361 CAAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFONALLVRYTKVPQVST 420  
DB 361 CAAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFONALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLNKGKVGSKCKKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRLNKGKVGSKCKKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKCCKADDKETCFABEGKCLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKCCKADDKETCFABEGKCLVAASQAALGL 585

RESULT 7

US-09-832-501-18  
; Sequence 18, Application US/09832501  
; Publication No. US20030199043A1  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J.  
; APPLICANT: Sleep, Darrell  
; APPLICANT: Turner, Andrew J.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF542  
; CURRENT APPLICATION NUMBER: US/09/832,501  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-832-501-18

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSVAHRFKDLGSENFKALVLIAPYQLOCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 1 DAHKSVAHRFKDLGSENFKALVLIAPYQLOCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
DB 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECCEKPLLEKSHCHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECCEKPLLEKSHCHCIAEVENDEMPA 300  
QY 301 DLPSSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360  
DB 301 DLPSSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360  
QY 361 CAAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFONALLVRYTKVPQVST 420  
DB 361 CAAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFEOQLGEYKFONALLVRYTKVPQVST 420  
QY 421 PTLVEVSRLNKGKVGSKCKKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRLNKGKVGSKCKKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVGHKPKAT 540  
QY 541 KEQLKAVMDDFAAFVEKCKCCKADDKETCFABEGKCLVAASQAALGL 585  
DB 541 KEQLKAVMDDFAAFVEKCKCCKADDKETCFABEGKCLVAASQAALGL 585

RESULT 8

US-09-833-118-18  
; Sequence 18, Application US/09833118  
; Publication No. US20030219875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF544  
; CURRENT APPLICATION NUMBER: US/09/833,118  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-833-118-18

Query Match 100.0%; Score 3103; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGSENFKALVLIAPYQLOCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 1 DAHKSVAHRFKDLGSENFKALVLIAPYQLOCCPFEDHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHNEETFLKKYLYEYIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180

```
Db 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLIP 180
QY 181 KLDELDRDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
Db 181 KLDELDRDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNQALLVRYTKVPQVST 420
Db 361 CAADAPHECYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNQALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
Db 481 LVNRRPCFSALVEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
```

## RESULT 9

```
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18
```

```
Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120
Db 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLIP 180
Db 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLIP 180
QY 181 KLDELDRDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
```

```
Db 181 KLDELDRDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNQALLVRYTKVPQVST 420
Db 361 CAADAPHECYAKVDFEFPKPLVEEPONLIKQNCLEPEQLGEYKFNQALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
Db 481 LVNRRPCFSALVEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
```

## RESULT 10

```
US-09-832-929-18
; Sequence 18, Application US/09832929
; Publication No. US20040171123A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF547
; CURRENT APPLICATION NUMBER: US/09/832,929
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-929-18
```

```
Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120
Db 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRIVRPEV 120
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLIP 180
Db 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAVTECCQAADKAACLIP 180
QY 181 KLDELDRDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
Db 181 KLDELDRDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
```

Db 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
QY 361 CAADPHCYAKVDFEFKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420  
Db 361 CAADPHCYAKVDFEFKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
Db 481 LVNRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585

## RESULT 11

US-10-153-604A-5  
; Sequence 5, Application US/10153604A  
; Publication NO. US20030143191A1  
; GENERAL INFORMATION:

; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,604A  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; TYPE: PRT  
; ORGANISM: Homo Sapiens

## US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
QY 121 DVWCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLPL 180  
Db 121 DVWCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLPL 180  
QY 181 KLDELURDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
Db 181 KLDELURDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
QY 361 CAADPHCYAKVDFEFKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420

Db 361 CAADPHCYAKVDFEFKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 421 PTLVEVSRNLGVSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
Db 481 LVNRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEEGKGLVAASQAALGL 585

## RESULT 12

US-10-319-263-1  
; Sequence 1, Application US/10319263  
; Publication NO. US20030180820A1  
; GENERAL INFORMATION:

; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/319,263  
; PRIOR FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens

## US-10-319-263-1

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
QY 121 DVWCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLPL 180  
Db 121 DVWCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLPL 180  
QY 181 KLDELURDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
Db 181 KLDELURDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVSKDVCCKNAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
 DB 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
 QY 421 PTLVEVSNLKGKVGSKCKHPKAPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCCTES 480  
 DB 421 PTLVEVSNLKGKVGSKCKHPKAPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCCTES 480  
 QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540  
 DB 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585  
 DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 13  
 US-10-319-263-2  
 ; Sequence 2, Application US/10319263  
 ; Publication No. US20030180820A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bar-Or M.D., David  
 ; APPLICANT: Lau Ph.D., Edward  
 ; APPLICANT: Winkler M.D., James V.  
 ; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
 ; FILE REFERENCE: ISCO07  
 ; CURRENT APPLICATION NUMBER: US/10/319,263  
 ; CURRENT FILING DATE: 2002-12-13  
 ; PRIOR APPLICATION NUMBER: 60/115,392  
 ; PRIOR FILING DATE: 1999-01-11  
 ; PRIOR APPLICATION NUMBER: 60/102,738  
 ; PRIOR FILING DATE: 1998-10-02  
 ; PRIOR APPLICATION NUMBER: 09/165,926  
 ; PRIOR FILING DATE: 1998-10-02  
 ; PRIOR APPLICATION NUMBER: 09/165,581  
 ; PRIOR FILING DATE: 1998-10-02  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: MOD RES  
 ; LOCATION: (1)..(585)  
 ; OTHER INFORMATION: ACETYLTATION  
 US-10-319-263-2

Query Match 100.0%; Score 3103; DB 14; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREPV 120  
 DB 61 NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREPV 120  
 QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLPL 180  
 DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLPL 180  
 QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
 DB 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLLEKC 360  
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLLEKC 360  
 QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
 DB 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420  
 QY 421 PTLVEVSNLKGKVGSKCKHPKAPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCCTES 480  
 DB 421 PTLVEVSNLKGKVGSKCKHPKAPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCCTES 480  
 QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540  
 DB 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585  
 DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 14  
 US-10-414-469-1  
 ; Sequence 1, Application US/10414469  
 ; Publication No. US20030190691A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bar-Or M.D., David  
 ; APPLICANT: Lau Ph.D., Edward  
 ; APPLICANT: Winkler M.D., James V.  
 ; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
 ; FILE REFERENCE: ISCO07  
 ; CURRENT APPLICATION NUMBER: US/10/414,469  
 ; CURRENT FILING DATE: 2003-04-15  
 ; PRIOR APPLICATION NUMBER: 09/806,247  
 ; PRIOR FILING DATE: 1999-10-01  
 ; PRIOR APPLICATION NUMBER: PCT/US99/22905  
 ; PRIOR FILING DATE: 1999-10-01  
 ; PRIOR APPLICATION NUMBER: 60/115,392  
 ; PRIOR FILING DATE: 1999-01-11  
 ; PRIOR APPLICATION NUMBER: 60/102,738  
 ; PRIOR FILING DATE: 1998-10-02  
 ; PRIOR APPLICATION NUMBER: 09/165,926  
 ; PRIOR FILING DATE: 1998-10-02  
 ; PRIOR APPLICATION NUMBER: 09/165,581  
 ; PRIOR FILING DATE: 1998-10-02  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-414-469-1

Query Match 100.0%; Score 3103; DB 14; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREPV 120  
 DB 61 NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREPV 120  
 QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLPL 180  
 DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLPL 180  
 QY 181 KLDELREDEGKASSAKORLKCASLOKFGBRAPKAWAVARLSORFPKAEFAEVSCLVTDLT 240

Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLT 240  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQQLGEYKFQNALLVRYTKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQQLGEYKFQNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585

RESULT 15  
US-10-414-469-2  
; Sequence 2, Application US/10414469  
; Publication No. US20030190691A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISCO07  
; CURRENT APPLICATION NUMBER: US/10/414,469  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: 09/806,247  
; PRIOR FILING DATE: 1999-10-01  
; PRIOR APPLICATION NUMBER: PCT/US99/22905  
; PRIOR FILING DATE: 1999-10-01  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (1)..(585)  
; OTHER INFORMATION: ACETYLYATION  
US-10-414-469-2

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAB 60  
Db 1 DAHSEVAHRFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAB 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAQBERNECFLOHKDDNPNLRLVRPEV 120

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAQBERNECFLOHKDDNPNLRLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKVLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKVLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180  
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLT 240  
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLT 240  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300  
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHGICIAEVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFLEYARRHPDYSVLLLRLLAKTYETTTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQQLGEYKFQNALLVRYTKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQQLGEYKFQNALLVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGVSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDDFAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585

RESULT 16  
US-10-413-831-1  
; Sequence 1, Application US/10413831  
; Publication No. US20030194813A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISCO07  
; CURRENT APPLICATION NUMBER: US/10/413,831  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: US/09/806,247  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-413-831-1

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAB 60

Db 1 DAHSEVAHRFKDLGEEFNFKALVLIATAQYLQOCFFEDHVKLVNEVTEFAKTCVADSAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Qy 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240  
Db 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240  
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTLEKC 360  
Qy 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420  
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 17

US-10-413-831-2  
; Sequence 1, Application US/10413831  
; Publication No. US20030194813A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/413,831  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: US/09/806,247  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(585)  
; OTHER INFORMATION: ACETYLYATION

US-10-413-831-2

Query Match 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 6.3e-251; Mismatches 0; Indels 0; Gaps 0;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHRFKDLGEEFNFKALVLIATAQYLQOCFFEDHVKLVNEVTEFAKTCVADSAE 60  
Db 1 DAHSEVAHRFKDLGEEFNFKALVLIATAQYLQOCFFEDHVKLVNEVTEFAKTCVADSAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Qy 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240  
Db 181 KLDELDRDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240  
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTLEKC 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLLRLAKTYETTLEKC 360  
Qy 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420  
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420  
Qy 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

RESULT 18

US-10-413-832-1  
; Sequence 1, Application US/10413832  
; Publication No. US20030215359A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/413,832  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: US/09/806,247  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-413-832-1

Query Match      100.0%; Score 3103; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHGICAEVNDMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHGICAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420
QY 421 PTLVEVSRLNKGKSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNKGKSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 19
US-10-413-832-2
; Sequence 2, Application US/10413832
; Publication No. US20030215359A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/413,832
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLATION
US-10-413-832-2

Query Match      100.0%; Score 3103; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAOYLQOCPEFHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHGICAEVNDMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHGICAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420
QY 421 PTLVEVSRLNKGKSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNKGKSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKGLVAASQAALGL 585

RESULT 20
US-10-414-386-1
; Sequence 1, Application US/10414386
; Publication No. US20030215952A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/414,386
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
```

```
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-386-1

Query Match      100.0%; Score 3103; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e-251;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFKDLGEENFKALVIAFAQYIQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHRFKDLGEENFKALVIAFAQYIQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120

Qy 121 DVMCTAFHDNEETFLKKVLYIARIHPYFYAPPELLFFAKRYKAAFTCCQADKAACLIP 180
Db 121 DVMCTAFHDNEETFLKKVLYIARIHPYFYAPPELLFFAKRYKAAFTCCQADKAACLIP 180

Qy 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPAEFAEVSKLVTDLTK 240

Qy 241 VHTCCHGDLLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCHGDLLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300

Qy 301 DLPSLAADPVSKDVKYAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360
Db 301 DLPSLAADPVSKDVKYAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360

Qy 361 CAADPHECYAKVDFEPKPLVEEPONLIKONCELFEOQLGEYKFQNALIVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEPKPLVEEPONLIKONCELFEOQLGEYKFQNALIVRYTKKVPQVST 420

Qy 421 PTLVEVSRNLGKVGSKCKHPKAMPKAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPKAMPKAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480

Qy 481 LVNRRPCPSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 481 LVNRRPCPSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540

Qy 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEGKKLVAASQAALGL 585
```

Search completed: October 13, 2005, 08:53:52  
Job time : 171 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:40:01 ; Search time 43 Seconds

(without alignments)  
1308.995 Million cell updates/sec

Title: US-10-816-042-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGEENFK.....TCFABEGKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop: 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	609	1 ABHUS	serum albumin prec
2	2942	94.8	600	2 A47391	serum albumin prec
3	2620	84.4	608	2 S57632	serum albumin prec
4	2475.5	79.8	607	1 ABHOS	serum albumin prec
5	2446.5	78.8	607	1 ABBO5	serum albumin prec
6	2432.5	78.4	607	1 ABHSH	serum albumin prec
7	2426	78.2	608	1 ABRTS	serum albumin prec
8	2411.5	77.7	605	1 ABPGS	serum albumin prec
9	2387	76.9	609	2 JCS838	albumin - Mongolia
10	1861	60.0	453	2 A05139	serum albumin - mo
11	1557.5	50.2	615	1 ABCHS	serum albumin prec
12	1253.5	40.4	609	2 JCA258	alpha-fetoprotein
13	1249.5	40.3	609	1 FPHU	alpha-fetoprotein
14	1242.5	40.0	609	1 FPGO	alpha-fetoprotein
15	1205	38.8	607	1 ABXL72	74K albumin prec
16	1181.5	38.1	265	2 I46986	albumin - dog (fra
17	1175.5	37.9	608	1 ABXL68	68K serum albumin
18	1084	34.9	605	1 FPMS	alpha-fetoprotein
19	1067	34.4	611	1 FPRT	alpha-fetoprotein
20	1055	34.0	599	1 A54906	afamin precursor
21	928.5	29.9	614	2 S59517	serum albumin prec
22	928	29.9	608	2 A53195	afamin precursor
23	747.5	24.1	608	1 ABONS1	serum albumin 1 pr
24	742.5	23.9	608	1 ABONS2	serum albumin 2 pr
25	699	22.5	382	2 A37253	serum albumin - bu
26	440.5	14.2	1423	1 S27941	serum albumin - se
27	386	12.4	474	1 VYHDD	vitamin D-binding
28	385	12.4	476	1 VYRTD	vitamin D-binding
29	372	12.0	472	1 A35327	vitamin D-binding

30	184	5.9	1819	2	A71928	cag island protein
31	184	5.9	1527	2	G64585	cag pathogenicity
32	141.5	4.6	1560	2	T30282	calcium-binding pr
33	137.5	4.4	1348	2	AG2558	hypothetical prote
34	134.5	4.3	1004	2	JC2221	major surface glyco
35	134	4.3	1083	2	JC2300	cell surface glyco
36	134	4.3	1780	2	T17272	hypothetical prote
37	133.5	4.3	1390	2	S51364	sperm tail-specifi
38	132.5	4.3	1070	2	T06733	kinesin homolog F2
39	131	4.2	1076	2	JC2217	major surface glyco
40	130	4.2	1175	2	PC3815	myosin heavy chain
41	129	4.2	1017	2	D34035	cell-cycle-depende
42	128.5	4.1	1051	2	JC4091	glycoprotein A - P
43	128	4.1	1005	2	A64465	hypothetical prote
44	128	4.1	3225	2	S52300	giantin - human
45	126.5	4.1	1927	2	A59236	embryonic muscle m
46	126.5	4.1	1972	1	A41604	myosin heavy chain
47	126.5	4.1	1974	2	T30010	hypothetical prote
48	126	4.1	1189	2	A54817	ATPase SC11, chrom
49	126	4.1	3259	1	A56539	giantin - human
50	124.5	4.0	1790	2	S67593	transport protein
51	124	4.0	2663	1	S28261	centromere protein
52	123.5	4.0	621	2	S10450	myosin heavy chain
53	123	4.0	1008	2	T10544	major surface glyco
54	123	4.0	1295	2	T24587	hypothetical prote
55	123	4.0	1538	2	T29095	cardiac muscle fac
56	123	4.0	1940	1	A24922	myosin heavy chain
57	122.5	3.9	344	2	S34153	mst101-1 protein -
58	122.5	3.9	1938	2	JC5421	smooth muscle myos
59	122.5	3.9	1972	2	JC5420	smooth muscle myos
60	122	3.9	779	2	T05990	hypothetical prote
61	122	3.9	2385	2	A34491	myosin heavy chain
62	122	3.9	2411	2	B32491	myosin heavy chain
63	121.5	3.9	932	2	S65214	probable alpha/gam
64	121.5	3.9	1280	2	A39117	170K lectin prec
65	121.5	3.9	1658	2	T42642	phosphoinositide 3
66	121	3.9	2253	2	T30336	nuclear/mitotic ap
67	120	3.9	1597	2	S68420	citron - mouse
68	119.5	3.9	946	2	S28061	SCP1 protein - rat
69	119.5	3.9	1056	1	G02157	kinesin-like spind
70	119	3.8	1002	2	T30546	major surface glyco
71	119	3.8	1175	2	C35815	myosin heavy chain
72	119	3.8	1201	2	A35815	myosin heavy chain
73	119	3.8	2297	2	AB2494	hypothetical prote
74	118.5	3.8	1010	2	AH2553	hypothetical prote
75	118.5	3.8	1128	2	G86266	hypothetical prote
76	118.5	3.8	1899	2	T32732	PAM C-terminal int
77	118.5	3.8	1919	2	T42098	PAM interacting pr
78	118	3.8	700	2	S67610	probable membrane
79	118	3.8	1201	2	B35815	myosin heavy chain
80	117.5	3.8	963	1	A41919	kinesin heavy chai
81	117	3.8	911	2	S51441	hypothetical prote
82	117	3.8	993	2	S49461	synaptonemal compl
83	116.5	3.8	734	2	T27055	hypothetical prote
84	116	3.7	1282	2	JE0120	glycoprotein A - m
85	116	3.7	1313	2	F96673	hypothetical prote
86	116	3.7	1937	2	T38055	myosin heavy chain
87	116	3.7	2007	1	B43402	myosin heavy chain
88	116	3.7	2168	2	T30171	ninein - mouse
89	115.5	3.7	955	4	C40045	probable transcrip
90	115.5	3.7	1388	2	T30335	KLP2 protein - Afr
91	115.5	3.7	1875	1	S38173	myosin-like protei
92	115	3.7	576	1	B35128	DNA repair and gen
93	115	3.7	1061	2	C88690	protein F41H10.4 [
94	115	3.7	1642	2	T08880	NMDA receptor-bind
95	115	3.7	1727	2	T50073	myosin-like coiled
96	115	3.7	1961	1	A61231	myosin heavy chain
97	115	3.7	1976	2	A52532	myosin heavy chain
98	114.5	3.7	765	2	F96558	hypothetical prote
99	114.5	3.7	1136	2	F96564	hypothetical prote
100	114.5	3.7	1940	2	A29320	myosin heavy chain

## ALIGNMENTS

## RESULT 1

ABHUS serum albumin precursor [validated] - human  
 N/Alternate names: preproalbumin  
 N/Contains: kinetensin  
 C/Species: Homo sapiens (man)  
 C/Date: 29-Jul-1981 #sequence\_revision 31-Jan-1997 #text change 09-Jul-2004  
 C/Accession: A93743; I3936; I59286; I59313; G01747; S55314; A91420; S06422; S36  
 R/Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seeburg  
 Nucleic Acids Res. 9, 6103-6114, 1981  
 A/Title: The sequence of human serum albumin cDNA and its expression in *Escherichia coli*  
 A/Reference number: A93743; MUID:82081882; PMID:6171778  
 A/Accession: A93743  
 A/Molecule type: mRNA  
 A/Residues: 1-419, 'K', 122-609 <LAW>  
 A/Cross-references: UNIPROT:P02768; EMBL:V00495; GB:J00078; GB:L00133; NID:G2  
 R/Dugaiczky, A.; Law, S.W.; Dennison, O.E.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
 A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.  
 A/Reference number: A93936; MUID:82105994; PMID:6275391  
 A/Accession: A93936  
 A/Molecule type: mRNA  
 A/Residues: 1-120, 'G', 122-609 <BUG>  
 A/Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590  
 R/Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
 J. Biol. Chem. 261, 3244-3251, 1986  
 A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and  
 A/Reference number: I39427; MUID:86140099; PMID:2419329  
 A/Accession: I39427  
 A/Status: translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-26 <URA>  
 A/Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173  
 R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
 A/Title: A nucleotide insertion and a frameshift cause analbuminemia in an Italian family.  
 A/Reference number: I59286; MUID:94181575; PMID:8134387  
 A/Accession: I59286  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 282-290, 'KSRFDLQ' <WAT>  
 A/Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033  
 R/Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
 R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,  
 Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
 A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-  
 A/Reference number: I59313; MUID:94294404; PMID:8022807  
 A/Accession: I59313  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 589-590, 'ALPRRVKNLLQVKLP' <MAD>  
 A/Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232  
 R/Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
 R/Menaya, J.; Parrilla, R.; Ayuso, M.S.  
 submitted to the EMBL Data Library, March 1995  
 A/Reference number: G08292  
 A/Accession: G08292  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-120, 'G', 122-455 <VEN>  
 A/Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431  
 R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
 Biochem. J. 308, 321-325, 1995  
 A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*  
 A/Reference number: S55314; MUID:9575251; PMID:7755581  
 A/Accession: S55314  
 A/Molecule type: protein  
 A/Residues: 19-27 <LED>  
 R/Meloun, B.; Moravsek, L.; Kostka, V.  
 FEBS Lett. 58, 134-137, 1975

A/Title: Complete amino acid sequence of human serum albumin.  
 A/Reference number: A91420; MUID:76187907; PMID:1225573  
 A/Accession: A91420  
 A/Molecule type: protein  
 A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-  
 R/Roehr, U.; Spitteller, G.; Tripiet, D.  
 Justus Liebig's Ann. Chem. 9, 881-884, 1988  
 A/Title: Isolation and structure elucidation of middle-molecular weight peptides from ur  
 A/Reference number: S06422  
 A/Note: this paper is in German, with an English abstract  
 A/Accession: S06422  
 A/Molecule type: protein  
 A/Residues: 25-48 <ROB>  
 R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
 Arch. Biochem. Biophys. 305, 595-599, 1993  
 A/Title: Mass spectrometric identification of modifications to human serum albumin treat  
 A/Reference number: S36882; MUID:93384321; PMID:8373198  
 A/Accession: S36882  
 A/Molecule type: protein  
 A/Residues: 45-67, 141-160; 311-337; 469-490; 570-581 <PIN>  
 R/Kausler, E.; Spitteller, G.  
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
 A/Title: Bruchtsuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol  
 A/Reference number: S17599; MUID:92126241; PMID:1772598  
 A/Accession: S17599  
 A/Molecule type: protein  
 A/Residues: 25-54, 354-357; 431-447 <KAU>  
 R/Note: 49-Leu was also found  
 R/Caraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
 J. Immunol. 143, 1680-1684, 1989  
 A/Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
 A/Reference number: A45800; MUID:89341406; PMID:2474609  
 A/Accession: A45800  
 A/Molecule type: protein  
 A/Residues: 166-173, 'L' <MOG>  
 R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa  
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
 A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre  
 A/Reference number: A03239; MUID:86242180; PMID:3087352  
 A/Accession: A03239  
 A/Molecule type: protein  
 A/Residues: 166-173, 'L' <MOG>  
 R/Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
 A/Title: Mutations in genetic variants of human serum albumin found in Italy.  
 A/Reference number: A38255; MUID:91062352; PMID:2247440  
 A/Accession: A38255  
 A/Molecule type: protein  
 A/Residues: 76-111 <GAL1>  
 A/Accession: B38255  
 A/Molecule type: protein  
 A/Residues: 82-105, 'K', 107-110 <GAL2>  
 A/Note: this variant is designated albumin Vibo Valentia  
 A/Accession: A38255  
 A/Molecule type: protein  
 A/Residues: 76-83, 'K', 85-106 <GAL3>  
 R/Note: this variant is designated albumin Torino  
 R/Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Fenni, R.  
 Eur. J. Biochem. 214, 437-444, 1993  
 A/Title: The structural characterization and bilirubin-binding properties of albumin Her  
 A/Reference number: S33298; MUID:93292504; PMID:8513793  
 A/Accession: S33298  
 A/Molecule type: protein  
 A/Residues: 255-263, 'E', 265-281 <MIN1>  
 R/Note: this variant is designated albumin Herborn  
 R/Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,  
 Biochim. Biophys. Acta 1119, 232-238, 1992  
 A/Title: Two albumins with identical electrophoretic mobility are produced by differ  
 A/Reference number: S21078; MUID:92190239; PMID:1347703  
 A/Accession: S21078  
 A/Molecule type: protein  
 A/Residues: 354-356, 'K', 358-378 <MIN2>  
 A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,

R:He, X.M.; Carter, D.C.  
 Nature 358, 209-215, 1992  
 A:Title: Atomic structure and chemistry of human serum albumin.  
 A:Reference number: A46756; MUID:92334427; PMID:1630489  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.  
 in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
 A:Reference number: A94442  
 A:Contents: annotation; three-dimensional structure and disulfide bonds  
 R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
 Collect. Czech. Chem. Commun. 42, 564-579, 1977  
 A:Title: Disulfide bonds in human serum albumin.  
 A:Reference number: A90930  
 A:Contents: annotation; disulfide bonds  
 R:Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A:Reference number: A90299; MUID:78186630; PMID:656055  
 A:Contents: annotation; bilirubin-binding site  
 R:Peters, T.; Reed, R.G.  
 in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20,  
 A:Title: Serum albumin: conformation and active sites.  
 A:Reference number: A94408  
 A:Contents: annotation; binding sites  
 R:Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
 A:Reference number: A90028; MUID:83279982; PMID:6192711  
 A:Contents: annotation; gene position  
 R:Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
 A:Reference number: A46755; MUID:76257808; PMID:955075  
 A:Contents: annotation  
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid  
 R:Bohney, J.P.; Fonda, M.L.; Felthoff, R.C.  
 FEBS Lett. 298, 266-268, 1992  
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphat  
 A:Reference number: A56294; MUID:92183881; PMID:1544460  
 A:Contents: annotation  
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p  
 atase activity  
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C:Comment: A large number of variants of human serum albumin have been described.  
 C:Genetics:  
 A:Gene: GDB:ALB  
 A:Cross-references: GDB:118990; OMIM:103600  
 A:Map position: 4q11-4q13  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status experimental <PRO>  
 F:25-609/Product: serum albumin #status experimental <MPT>  
 F:29-202/Domain: serum albumin repeat homology <SAI>  
 F:166-174/Product: kinetinsin #status experimental <KIP>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:27/Binding site: copper (His) #status predicted  
 F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 9,3e-199;  
 Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHKEVAHRFKDLGEENFKALVLIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60  
 DB 25 DAHKEVAHRFKDLGEENFKALVLIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQPERNECFLOHKDDNPPLRLVRPEV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQPERNECFLOHKDDNPPLRLVRPEV 144

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
 DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204  
 QY 181 KLDELURDSEKASSAKORLKCASLOKFGSRAFKAWAVARLSORFPAEFAEVSKLVTDLTK 240  
 DB 205 KLDELURDSEKASSAKORLKCASLOKFGSRAFKAWAVARLSORFPAEFAEVSKLVTDLTK 264  
 QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHKSHCIAEVENDEMPA 300  
 DB 265 VHTECHGDLLECADRADLAKYICENQDSISSKLKECEKPLLEKSHKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYVEYARRHPDYVSVLLRLRAKYETTTLEKC 360  
 DB 325 DLPSLAADFVESKOVCKNYAEAKOVFLGMFLYVEYARRHPDYVSVLLRLRAKYETTTLEKC 384  
 QY 361 CAAADPHCEYAKVDFEFKPLVEEPONLIKONCELEFQOLGEYKFONALLVRYTKVPOVST 420  
 DB 385 CAAADPHCEYAKVDFEFKPLVEEPONLIKONCELEFQOLGEYKFONALLVRYTKVPOVST 444  
 QY 421 PTLVEVSRLNKGVSCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEVSRLNKGVSCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHKPKAT 540  
 DB 505 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKGHKPKAT 564  
 QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKLVVAASQAALGL 585  
 DB 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEEGKLVVAASQAALGL 609

## RESULT 2

A47391  
 serum albumin precursor - rhesus macaque  
 C:Species: Macaca mulatta (rhesus macaque)  
 C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C:Accession: A47391  
 R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
 A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilir  
 A:Reference number: A47391; MUID:93211971; PMID:8460152  
 A:Contents: B/B homozygote  
 A:Accession: A47391  
 A:Status: preliminary  
 A:Molecule type: mRNA; protein  
 A:Residues: 1-600 <WAT>  
 A:Cross-references: UNIPROT:Q28522; GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G34229  
 A:Experimental source: liver  
 A:Note: sequence extracted from NCBI backbone (NCBI:128280, NCBIP:128281)  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:21-194/Domain: serum albumin repeat homology <SA1>  
 F:213-386/Domain: serum albumin repeat homology <SA2>  
 F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 94.8%; Score 2942; DB 2; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 4.6e-188;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60  
 DB 17 DTHKEVAHRFKDLGEENFKALVLIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQPERNECFLOHKDDNPPLRLVRPEV 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQPERNECFLOHKDDNPPLRLVRPEV 136  
 QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180  
 DB 137 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 196



Db 204 KLDALKERILLSSAKERLKCSEFQNGPERAVKAVSVARLSQKPKADFAEVSKIYDITLTK 263  
QY 241 VTECHGDLLECADRADLAKYICENQDSISSKLKCECEKPLLEKSHCIAEVENDEMPA 300  
Db 264 VHKECHGDLLECADRADLAKYICHQDSISGKLKACCDKPLQKSHCIAEVEKDDLPS 323  
QY 301 DLPSLAADPVESKDVCKYAEAKDVFLGMLFYEARHDPDYSVLLRLAKTYETLEKC 360  
Db 324 DLPALAADPAEDKCKYKADKDVFLGTLFYESRRHDPDYSVLLRLAKTYETLEKC 383  
QY 361 CAAADPHCYAKVDFDEKFLVEEPONLIIKONCEPEQGEYKFQNALIVRYTKVPQVST 420  
Db 384 CAEADPPACYRIVFDOFTFLVEEPSKLVKNCNDLPEEVEGYDFQNALIVRYTKAPQVST 443  
QY 421 PTLVEVRNLGVKGVKCKKHPAKRMPCAEADYLSVVLNOLCVLHKTPTVSDRVTKCCTES 480  
Db 444 PTLVEIGRTLGVKGVKCKKHPAKRMPCAEADYLSVVLNOLCVLHKTPTVSDRVTKCCTES 503  
QY 481 LVNRPRCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540  
Db 504 LAERPRCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 563  
QY 541 KQQLKAVMDDFAAFYKCKCKADDKETCFPAEKGKLVAAASQAL 583  
Db 564 KQQLKAVMDDFAAFYKCKCKADDKETCFPAEKGKLVAAASQAL 606

RESULT 5  
ABBS  
serum albumin precursor [validated] - bovine  
N:Alternate names: 67K protein; prealbumin  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 09-Jul-2004  
C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94551  
R:Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.  
submitted to the EMBL Data Library, August 1991  
A:Description: Bovine serum albumin: cDNA sequence and expression.  
A:Reference number: A38885  
A:Accession: A38885  
A:Molecule type: mRNA  
A:Residues: 1-607 <HOL>  
A:Cross-references: UNIPROT:P04277; EMBL:M73215  
R:Hiroyama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.  
Biochem. Biophys. Res. Commun. 173, 639-646, 1990  
A:Title: Rapid confirmation and revision of the primary structure of bovine serum albumin  
A:Reference number: A36401; MUID:91083649; PMID:2260975  
A:Accession: A36401  
A:Molecule type: protein  
A:Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>  
R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.  
Eur. J. Biochem. 98, 477-485, 1979  
A:Title: Biosynthesis of bovine plasma proteins in a cell-free system.  
A:Reference number: A91258; MUID:80024278; PMID:488109  
A:Accession: A91258  
A:Molecule type: protein  
A:Residues: 1-32 <MAG>  
R:Hsieh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-8, 1988  
A:Title: Electrophoresis onto glass-fiber filter from an analytical isoelectrofocusing gel  
A:Reference number: A60808; MUID:88267456; PMID:3389500  
A:Accession: B60808  
A:Molecule type: protein  
A:Residues: 25-41 <HSI>  
R:Strawich, E.; Glimcher, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A:Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin  
A:Reference number: S10780; MUID:90336641; PMID:2379503  
A:Accession: S10780  
A:Molecule type: protein  
A:Residues: 25-41, 'H', 43-57, 59-64 <STR>  
R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases

A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: D45800  
A:Molecule type: protein  
A:Residues: 163-172 <CAR>  
R:Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A:Title: Structure of a biologically active neurotensin-related peptide obtained from pm  
A:Reference number: A26693; MUID:87194805; PMID:2437111  
A:Accession: A26693  
A:Molecule type: protein  
A:Residues: 165-172, 'L', 'CA2'  
R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A:Title: Sequence of residues 400-403 of bovine serum albumin.  
A:Reference number: A90309; MUID:82023364; PMID:7283978  
A:Accession: A90309  
A:Molecule type: protein  
A:Residues: 402-433 <REE>  
R:Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A:Title: Structure of bovine serum albumin.  
A:Reference number: A91458  
A:Accession: A91458  
A:Molecule type: protein  
A:Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'P'  
submitted to the Atlas, April 1975  
A:Reference number: A94551  
A:Accession: A94551  
A:Molecule type: protein  
A:Residues: 190-195 <BR2>  
R:Brown, J.R.  
Fed. Proc. 33, 1389, 1974  
A:Reference number: A91457  
A:Contents: annotation; disulfide bonds  
R:Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC  
A:Reference number: S55232; MUID:95031935; PMID:7945219  
A:Accession: S55232  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 529-536; 569-572 <WER>  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; copper binding; duplication; plasma  
F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-24/Domain: propeptide #status experimental <PRO>  
F:25-607/Product: serum albumin #status experimental <MPT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,

Query Match 78.8%; Score 2446.5; DB 1; Length 607;  
Best Local Similarity 75.6%; Pred. No. 4.1e-155;  
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;  
QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 25 DTHKSEIARRPDKJGSEQFKGLVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADES 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRVLPRPV 120  
Db 85 GCEKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRVLPRPV 143  
QY 121 DMCTAFHDNDETFKVKLYEIARRHPYFYAPPELLFFAKRYKAAATECCQAADKACLLP 180  
Db 144 NTLCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRVLPRPV 203  
QY 181 KLDLDRDSEKASSAKORLKCKASLOKFGERAFKAMAVARLSQKPKADFAEVSKIYDITLTK 240  
Db 204 KIETMRKVLASSAKQRURCASIQKFGERALKAMAVARLSQKPKADFAEVSKIYDITLTK 263







**Qy**      541 KEQLKAVMDDFAAFVEKCKADDKETCAAEKGKLV 576  
              :|||: |: |||||:||| |---| ---|  
**Dd**      562 EEOGLTVLGNFAFVKOKCAAOPHEACFAVEGPKVF 597

## RESULT 9

JC5838  
albumin - Mongolian jird  
C:Species: Meriones unguiculatus (Mongolian jird)  
C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JC5838  
R:Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.  
DNA Res. 4, 351-354, 1997  
A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the  
A:Reference number: JC5838; MUID:98116663; PMID:9455485  
A:Accession: JC5838  
A:Molecule type: mRNA  
A:Residues: 1-609 <YOS>  
A:Cross-references: UNIPROT:O35090; DDBJ:AB006197; NID:g2317277; PIDN:BAA21765.1; PID:g2317277  
A:Experimental source: liver  
C:Superfamily: serum albumin; serum albumin repeat homology  
F:222-395/Domain: serum albumin repeat homology <SA2>

RESULT 10

A05139  
serum albumin - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 05-Jun-1987 #sequence revision 17-Mar-2000 #text change 09-Jul-2004







QY 417 QVSTPTLVEVSRNLGKVGSKCKHPKAMPKCAEDYLSVVLNOLCVLHKTPTVSDRVTKC 476  
 Db 441 QLTSSLMATRKMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEHMTVPNPGVQC 500  
 QY 477 CTESLVNRRPCFSALEVDEYVYKFEFNAETFTFHADICTLSEKEROIKKQATLVELVKHK 536  
 Db 501 CTSSYANRRPCFSLVVDYVYPPAFSDDKFIHKLCOAQGVALQTMKQEFNLINLVKQK 560  
 QY 537 PKATKEQLKAVMDDFAAFEVCKCKADKCTCFABEKGKLVAAASQAALGL 585  
 Db 561 PQITEQLEAVIADFSGLLEKCCQGEQEVCFABEKGKLSKTRTALGV 609  
 RESULT 14  
 FGQ  
 alpha-fetoprotein precursor - gorilla  
 C:Species: Gorilla gorilla (gorilla)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A37970  
 R:Ryan, S.C.; Zielinski, R.; Dugaiczkyk, A.  
 Genomics 9, 60-72, 1991  
 A:Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates.  
 A:Reference number: A37970; MUID:91169517; PMID:1706310  
 A:Accession: A37970  
 A:Molecule type: DNA  
 A:Residues: 1-609 <RYA>  
 A:Cross-references: UNIPROT:P28050; GB:M38272; NID:g817963; PID:AAA73520.1; PID:g177041  
 C:Genetics:  
 A:Map position: 4q11-12  
 A:Supernotes: 29/1; 46/2; 90/3; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551/3  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-609/Product: alpha-fetoprotein #status predicted <SA1>  
 F:29-202/Domain: serum albumin repeat homology <SA1>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:22/Binding site: copper (His) #status predicted  
 F:99-114,113-124,148-193,192-224-270,269-277,289-303,302-313,384-393,416-462,461-472  
 F:249/Binding site: bilirubin (Lys) #status predicted  
 F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.0%; Score 1242.5; DB 1; Length 609;  
 Best Local Similarity 39.8%; Pred. NO. 4.5e-75;  
 Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;  
 QY 3 HKSE-----VAHRFKDLGGENFKALVLIAPAOYLQCCPPEDHVKLVNVEVTEFAKTCVADE 57  
 Db 22 HRNEYGIASILDYQCTAISLADLATIFPAQVQATYKESKMWKDALTAIEKPTGDE 81  
 QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNP-NLPLRV 116  
 Db 82 QSAGCLENLQAPLEELCHEKEILEKYG-LSDCSSQSEEGHNCFLAHKPTPASIFLQ 140  
 QY 117 REVDVMTAFHDNETFLKLYIETARHPYFPAPELLFFAKYKAAATECCQADKAA 176  
 Db 141 VPEVPTSCAYEDRETFMKNFYEIARRHPFLYAPTILLAAARYDKIIPSCCKAENAYE 200  
 QY 177 CLLPKLDELDEKAGKASQRLKASQKQFGERAFKAWAVARLSQRFPAEFAEYKLVY 236  
 Db 201 CFQTKAATYTKELRESLLNQACAKWKQFGRTFOAITVTKLSQKFTKVNTEIQKLV 260  
 QY 237 DLTQVTECHGDLLECADRADLAKYICENQDISSKLKECECKPFLLEKSHCIAEVEND 296  
 Db 261 DVAAVHEHCGRGDLVLDQGEKIMSYICSQQDTLSNKITCECKLTTLERGQCIIHAEND 320  
 QY 297 ENPADLPSLAADPVESKDVCKNYAEAKDVFLGNFLYEARHPDYSVLLLRKITYET 356  
 Db 321 EKPEGLSPNLNRLGDRDFNQFSSGKNIPLASFVHEYSRRHPQLAVSVILRVAKGYQL 380  
 QY 357 LEKCCAAADPHCEYAKVDFEFLVPEPQNLKONCELFQELGEYKFNALLVRYTKVP 416  
 Db 381 LEKCFOTENPLEQDQGBEELQYIOESQALAKSCGLFQKLGYYLQNAFLVAYTKCAP 440

QY 417 QVSTPTLVEVSRNLGKVGSKCKHPKAMPKCAEDYLSVVLNOLCVLHKTPTVSDRVTKC 476  
 Db 441 QLTSSLMATRKMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEHMTVPNPGVQC 500  
 QY 477 CTESLVNRRPCFSALEVDEYVYKFEFNAETFTFHADICTLSEKEROIKKQATLVELVKHK 536  
 Db 501 CTSSYANRRPCFSLVVDYVYPPAFSDDKFIHKLCOAQGVALQTMKQEFNLINLVKQK 560  
 QY 537 PKATKEQLKAVMDDFAAFEVCKCKADKCTCFABEKGKLVAAASQAALGL 585  
 Db 561 PQITEQLEAVIADFSGLLEKCCQGEQEVCFABEKGKLSKTRTALGV 609  
 RESULT 15  
 ABXL72  
 74K albumin precursor - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: B41682; S02693; A05288  
 R:Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.  
 Mol. Endocrinol. 3, 464-473, 1989  
 A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.  
 A:Reference number: A41682; MUID:89313788; PMID:2747653  
 A:Accession: B41682  
 A:Molecule type: mRNA  
 A:Residues: 3-607 <MOS>  
 A:Cross-references: UNIPROT:P14872; GB:M21442; NID:G213930; PID:AAA49637.1; PID:G21393  
 R:Schorpp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U.  
 J. Mol. Biol. 199, 83-93, 1988  
 A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Del  
 A:Reference number: S02692; MUID:88172470; PMID:2451026  
 A:Accession: S02693  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-48 <SCH>  
 A:Cross-references: EMBL:226826  
 R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.  
 Eur. J. Biochem. 146, 489-496, 1985  
 A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilization  
 A:Reference number: A05288; MUID:85126974; PMID:3971963  
 A:Accession: A05288  
 A:Molecule type: mRNA  
 A:Residues: 459-502, 'L', 504-557 <WOL>  
 A:Cross-references: GB:M28276  
 A:Note: the authors translated the codon TAT for residue 63 as Thr  
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thy  
 C:Genetics:  
 A:Introns: 27/1  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status predicted <PRO>  
 F:25-607/Product: 74K serum albumin #status predicted <SA1>  
 F:32-201/Domain: serum albumin repeat homology <SA2>  
 F:202-393/Domain: serum albumin repeat homology <SA2>  
 F:412-591/Domain: serum albumin repeat homology <SA3>  
 F:30/Binding site: copper (His) #status predicted  
 F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392  
 F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 38.8%; Score 1205; DB 1; Length 607;  
 Best Local Similarity 39.3%; Pred. No. 1.4e-72;  
 Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;  
 QY 3 HKSEVAHRFKDLGGENFKALVLIAPAOYLQCCPPEDHVKLVNVEVTEFAKTCVADEAENC 62  
 Db 30 HHKHIADVYALTERTFKGLTLAIVSQNLKQCSLELSKLVNEINDFAKSCINDKTPE-C 88  
 QY 63 DKSHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNP-LPLRVREVDV 122

```
Db 89 EKPVGTLFFDKLCADPAVGAVGVNYEWSKECCAKQDPERACQFAKRAHDHEHT---SIKPEEE 145
Qy 123 MCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLPAKEYKAAFTCCCAADKAACLPLKL 182
Db 146 TCKLLKEHPDLLSAFIEEARNHPDLPPYPAVLATKQYHKLAEHCCBEEDKEKCFSEKM 205
Qy 183 DELRDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFPKAFBVSKLVDLT/KVH 242
Db 206 KQLMKQSHSIEDKQHHFCWILNDNPFKVLKALNARVSHRYPKAEFFKLHNFTBEVTHFI 265
Qy 243 TECCHGDLLECADRADLAKYICENQDSISKKLKECCCKPLEKSHCIAEVENDEMPADL 302
Db 266 KDCCHDDMFECWTERLELTHECTQHKLBSKKLECCNIPLLERTYICIVLTLENDVPAEL 325
Qy 303 PSLAADFVESKDVCNKYAEAKDVLGMLFELYEARHPDYSVVLRLAKTYTTLEKCCA 362
Db 326 SQPITFETDHPVCEKYAENNEVFGRYLHVAVRSHQBSLSEQFLQSAYKESLNLKCKC 385
Qy 363 AADPHCEYAKVDFEFPKLVPEEPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVSTPT 422
Db 386 TDNPECYKDGADRPMNAEKERFAVLKQNCDLIHGEGYLFENELLIRYTKMPQVSDET 445
Qy 423 LVEVSRNLKGVSKCKHPEAKRMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCTESLV 482
Db 446 LIGIAHQWADIGEHCACVPEQNMPCAEGBDILTILIGKMCERQKKTFINNHVACCTDSYS 505
Qy 483 NRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKQATVLELVKHKPKATKE 542
Db 506 GMRSCFTALGPDDEDYVPPVTDTHFDKDKICTANDKEKHQIKQKFLKLIKVPSPKLEKN 565
Qy 543 OLKAWMDFAAFVEKCKKADKQKTCFAEEGKKLVAASQ 580
Db 566 HIDECSAEFLKXWQKCCCTADEHQPCFDTKPKVLIEHQ 603

RESULT 16
I46986
A:Species: Canis lupus familiaris (dog)
C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999
C:Accession: I46986
R:Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner
J. Allergy Clin. Immunol. 93, 614-627, 1994
A:Title: Molecular characterization of dog albumin as a cross-reactive allergen.
A:Reference number: I46986; MUID:94201492; PMID:7512102
A:Accession: I46986
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-265 <SPI>
A:Cross-references: GB:S72946; NID:G633937; PIDN:AAB30434.1; PID:G633938
C:Superfamily: serum albumin; serum albumin repeat homology
F:7-180/Domain: serum albumin repeat homology <SA2>

Query Match 38.1%; Score 1181.5; DB 2; Length 265;
Best Local Similarity 82.6%; Pred. No. 1.9e-71;
Matches 218; Conservative 23; Mismatches 22; Indels 1; Gaps 1;
Qy 192 SSAKORLKCSAQKQGERAFKAWAVARLSQRPFPKAFBVSKLVDLT/KVHTECHGDL 251
Db 2 SSAKERFKCSAQKQGERAFKAWAVARLSQRPFPKADFAEISKVVDLT/KVHKECHGDL 61
Qy 252 ECADDRADLAKYICENQDSISKKLKECCCKPLEKSHCIAEVENDEMPADLPSLAADFE 311
Db 62 ECADDRADLAKYICENQDSISKKLKECCCKPVLEKSOCLAEVERDELPGDLPSLAADFE 121
Qy 312 SKDVCNKYAEAKDVLGMLFELYEARHPDYSVVLRLAKTYTTLEKCCAADPHCEYA 371
Db 122 DKEVCNKYQEAQKDVFLGTFLVEYSRRHPEYSVSLRLAKEATLEKCCATDDPPTCYA 181
Qy 372 KVFDEFPKLVPEEPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVSTPTL-VEVSRNL 430
Db 182 KYLDEFPKLVPEEPQNLVNTCNELFEKGEYGFQNALLVRYTKAPQVSTPTLVVEVSRKL 241
```

```
Qy 431 GKVSKCKKHPKAEAKRMPCAEYDLS 454
Db 242 GKVGTKCKKPESEKSCADDPLS 265

RESULT 17
ABXL68
A:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A41682; S02692
R:Moskatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
e during development.
A:Reference number: A41682; MUID:89313788; PMID:2747653
A:Accession: A41682
A:Molecule type: mRNA
A:Residues: 1-608 <MOS>
A:Cross-references: UNIPROT:P08759; GB:M18350
R:Schorpp, M.; Doeberling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Dele
A:Reference number: S02692; MUID:88172470; PMID:2451026
A:Accession: S02692
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:Z26825
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thyre
C:Genetics:
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-608/Product: 68K serum albumin #status predicted <MAT>
F:32-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89,102-118,117-128,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,
Query Match 37.9%; Score 1175.5; DB 1; Length 608;
Best Local Similarity 38.3%; Pred. No. 1.3e-70;
Matches 222; Conservative 111; Mismatches 240; Indels 7; Gaps 2;
Qy 3 HKSEVAHRFKDLGEENFKALVLIATAQYVLOQCPPEHDHVKLNVETFEAKTCVADESABNC 62
Db 30 HHKHIADWYNLLTERTFKGLTLAIVSQNLQKCSLEELSKLVNEINDFAKSGTGNDKTPEC 89
Qy 63 DKSUHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHK--DDNPMLPRLVRPEV 120
Db 90 EKPIGTFLFYDKLCAADPKVGVNYSWEKSCSKQDPERAQCFRAHRVFEHNP-----VRPKP 144
Qy 121 DVMCTAHDEETELKKYVLEIARRHPYFYAPPELLFAKRYKAAFTCCCAADKAACLPL 180
Db 145 EETCALFKHPDDLLSAFIEEARNHPDLPPYPAVLTLTQQYKLVHECCBEEDKDKCPAE 204
Qy 181 KLDELIRDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFPKAFBVSKLVDLT/K 240
Db 205 KMKELMKHSHSIEDKQHHFCWILNDNPFKVLKALNARVSHRYPKAEFFKLHNFTBEVTH 264
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISKKLKECCCKPLEKSHCIAEVENDEMPA 300
Db 265 FIKDCCHGDMFECWTERLELTHECTQHKLBSKKLECCNIPLLERTYICIVLTLENDVPA 324
Qy 301 DLPSLAADFVESKDVCNKYAEAKDVLGMLFELYEARHPDYSVVLRLAKTYTTLEKCC 360
Db 325 ELSRPTFETDHPVCEKYAENKESFLERSIPWQSQETPELSEQFLQSAKEYESLNLKC 384
Qy 361 CAAADPHCEYAKVDFEFPKLVPEEPQNLIKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420
```



A:Reference number: I53048; MUID:92119318; PMID:1722723  
A:Accession: I53048  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 7-52 <RES>  
A:Cross-references: GB:M18351; NID:g202783; PIDN:AAA68903.1; PID:g202784  
R:Nahon, J.  
J. Biol. Chem. 262, 12479-12487, 1987  
A:Title: The rat alpha-fetoprotein and albumin genes: Transcriptional control and comparison with the human genes  
A:Reference number: I55253; MUID:87308273; PMID:2442163  
A:Accession: I55253  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 7-19 <RES>  
A:Cross-references: GB:J02816; NID:g202778; PIDN:AAA40695.1; PID:g202779  
R:Liao, W.S.L.; Hamilton, R.W.; Taylor, J.M.  
J. Biol. Chem. 255, 8046-8049, 1980  
A:Title: Amino acid sequence homology between rat alpha-fetoprotein and albumin at the C-terminus  
A:Reference number: I55217; MUID:81006806; PMID:6157681  
A:Accession: I55217  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 'VG', 558-611 <RES>  
A:Cross-references: GB:J000694; NID:g202776; PIDN:AAA40694.1; PID:g202777  
R:Innis, M.A.; Miller, D.L.  
J. Biol. Chem. 255, 8994-8996, 1980  
A:Title: Alpha-fetoprotein gene expression. Partial DNA sequence and COOH-terminal homology  
A:Reference number: I55218; MUID:81006964; PMID:6157690  
A:Accession: I55218  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 476-525, 'L', 527-529, 'ATNSSTRN', 539-546, 'P', 548-597, 'VQVDFQ', 605-611 <RES>  
A:Cross-references: GB:J000693; NID:g56141; PIDN:CAA24546.1; PID:g56142  
A:Accession: I69969  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 476-525, 'L', 527-529, 'ATNSSTRN', 539-546, 'P', 548-597, 'VQVDFQ', 605-611 <RES>  
R:Chevrette, M.; Guertin, M.; Turcotte, B.; Belanger, L.  
Nucleic Acids Res. 15, 1338-1339, 1987  
A:Title: The rat alpha 1-fetoprotein gene: characterization of the 5'-flanking region and the 3'-untranslated region  
A:Reference number: I58260; MUID:87146445; PMID:2434929  
A:Accession: I58260  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-34 <RES>  
A:Cross-references: EMBL:X05093; NID:g55610; PIDN:CAA28744.1; PID:g55611  
A:Experimental source: Sprague-Dawley  
C:Comment: The rat AFP strongly binds estrogens.  
C:Genetics:  
A:Introns: 35/1  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: duplication; embryo; fetus; globulin; glycoprotein; metal binding; plasma  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-611/Product: alpha-fetoprotein #status predicted <MAT>  
F:35-204/Domain: serum albumin repeat homology <SA1>  
F:223-396/Domain: serum albumin repeat homology <SA2>  
F:287-611/Product: alpha-fetoprotein variant #status predicted <MVT>  
F:415-594/Domain: serum albumin repeat homology <SA3>  
F:28/Binding site: copper (His) #status predicted  
F:101-116, 115-126, 194-203, 226-272, 271-279, 291-305, 304-315, 386-395, 418-464, 463-474  
F:253/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 34.4%; Score 1067; DB 1; Length 611;  
Best Local Similarity 34.8%; Pred. No. 2.le-63;  
Matches 199; Conservative 126; Mismatches 239; Indels 8; Gaps 4;

Qy 16 EENFKALVLAFAYLOQCPEFHVKL-VNEVTFEAKTCVADSEACNDKSLHTLFGDKL 74  
Db 46 EKXMFNVATIVAGFVQ-----DATKAEVKNMSDALLAANKENTGDCLENQLSVFLDEI 100  
Qy 75 CTVATLRETYGEMADCCAKQEPERNECFLOHKDNP-NLPRLVRPEVDVVMCTAFHNEET 133



**THIS PAGE BLANK (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 08:39:21 ; Search time 178 Seconds  
(without alignments)  
1682.957 Million cell updates/sec

Title: US-10-816-042-18  
Perfect score: 3103  
Sequence: 1 DAHKEVAHRFKDLGSENFK.....TCFAEKGKLVAAASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	3103	100.0	609	1	ALBU HUMAN
2	3079	99.2	609	2	Q68DN5
3	2942	94.8	600	1	ALBU MACMU
4	2620	84.4	608	1	ALBU FELCA
5	2614	84.2	584	2	Q7YSG3
6	2562	82.6	608	1	ALBU CANFA
7	2504	80.7	608	2	Q9SVB7
8	2475.5	79.8	607	1	ALBU HORSE
9	2457	79.2	608	1	ALBU RABIT
10	2450.5	79.0	607	1	ALBU BOVIN
11	2432.5	78.4	607	1	ALBU SHEEP
12	2426	78.2	608	1	ALBU RAT
13	2411.5	77.7	605	1	ALBU_PIG
14	2404.5	77.5	607	2	Q68NH7
15	2387	76.9	608	2	Q68DN9
16	2379.5	76.7	583	1	ALBU MERUN
17	2378	76.6	608	1	ALBU MOUSE
18	2374	76.5	608	2	Q8C7H3
19	2336	75.3	576	2	Q8C7C7
20	2336	75.3	417	2	Q8CYGO
21	1991	64.2	21	1991	64.2
22	1865.5	60.1	396	2	Q8IUK7
23	1557.5	50.2	615	1	ALBU CHICK
24	1295.5	41.7	527	2	Q8JIA9
25	1253.5	40.4	609	1	FETA_PANTR
26	1249.5	40.3	609	1	FETA_HUMAN
27	1242.5	40.0	609	1	FETA_GORGO
28	1242	40.0	609	2	Q8MJU5
29	1242	40.0	626	2	Q8UW05
30	1218.5	39.3	610	2	Q8MJ76
31	1205	38.8	607	1	ALB2_XENLA

32	1200	38.7	609	1	FETA_HORSE
33	1199	38.6	607	2	Q642P7
34	1183.5	38.1	608	2	Q7T5F3
35	1164.5	37.5	606	1	ALB1_XENLA
36	1087	35.0	624	2	Q8UW06
37	1084	34.9	605	1	FETA_MOUSE
38	1082	34.9	605	2	Q8BK65
39	1076	34.7	605	2	Q8BK56
40	1067	34.4	611	1	FETA_RAT
41	1055	34.0	599	1	AFAM_HUMAN
42	1045	33.7	400	2	Q8JIA7
43	955	30.8	603	2	Q9YGH6
44	944	30.4	611	1	AFAM_MOUSE
45	928.5	29.9	614	2	Q9I134
46	928	29.9	608	2	AFAM_RAT
47	888	28.6	406	2	Q8JIA8
48	771	24.8	167	2	Q6UXK4
49	747.5	24.1	608	1	ALB1_SALSA
50	742.5	23.9	608	1	ALB2_SALSA
51	739	23.8	205	2	Q8CG74
52	713	23.0	417	2	Q8ROJ9
53	699	22.5	382	1	ALBU_RANCA
54	440.5	14.2	1423	1	ALBU_PETMA
55	386	12.4	474	1	VTDB_HUMAN
56	385	12.4	476	2	Q68FY4
57	381	12.3	476	1	VTDB_RAT
58	378	12.2	476	1	VTDB_RABIT
59	376.5	12.1	484	2	Q9W6F5
60	372	12.0	476	1	VTDB_MOUSE
61	365	11.8	551	2	Q42279
62	331	10.7	122	2	Q90WZ8
63	290	9.3	123	2	Q90WZ6
64	284	9.2	464	2	Q6DGV8
65	203	6.5	1894	2	Q75XB6
66	200	6.4	1927	2	Q75WX9
67	198.5	6.4	1820	2	Q75XM5
68	196	6.3	807	2	Q83ZS7
69	195	6.3	1810	2	Q75X88
70	195	6.3	2002	2	Q75XH1
71	194.5	6.2	1813	2	Q75XE3
72	192.5	6.2	1797	2	Q75X61
73	190.5	6.1	1795	2	Q75XJ8
74	188	6.1	1723	2	Q9JMX8
75	186.5	6.0	1796	2	Q75XQ2
76	184	5.9	1819	2	Q92LV0
77	184	5.9	1927	2	Q25262
78	183	5.9	1802	2	Q75X05
79	177.5	5.7	1768	2	Q75X33
80	162.5	5.2	44	2	Q95MC2
81	162.5	5.2	680	2	Q9V6S8
82	161	5.2	1079	2	Q96V11
83	156	5.0	1026	2	Q74669
84	154.5	5.0	3843	2	Q9USD0
85	153.5	4.9	661	2	Q8MS79
86	153.5	4.9	3843	2	Q9VU94
87	153	4.9	62	2	Q8M11
88	152.5	4.9	810	2	Q64JVA
89	151.5	4.9	1605	1	RRB1_MOUSE
90	150.5	4.9	1077	2	Q6AHS1
91	150	4.8	1065	2	Q01828
92	149	4.8	40	2	Q9TRA5
93	149	4.8	897	2	Q64JW5
94	148	4.8	804	2	Q62PS0
95	148	4.8	1028	2	Q74668
96	145	4.7	1825	2	Q7RIT2
97	144.5	4.7	8407	2	Q7ETM4
98	144.5	4.7	8797	1	SNE1_HUMAN
99	143	4.6	1069	2	Q6AHV5
100	142.5	4.6	218	2	Q66RQ0

ALIGNMENTS

## RESULT 1

ALBU HUMAN  
ID \_ALBU\_HUMAN STANDARD; PRT; 609 AA.  
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJ20;  
DT 21-JUL-1986 (Rel. 01, Created)  
DD 01-APR-1990 (Rel. 14, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Serum albumin precursor.  
GN Name=ALB;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196112; PubMed=3009475;  
RA Minghetti P.P., Ruffner D.B., Kuang W.J., Dennison O.E., Hawkins J.W.,  
RT Beattie W.G., Dugaiczak A.;  
RT "Molecular structure of the human albumin gene is revealed by  
RT nucleotide sequence within q11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
RX MEDLINE=82081882; PubMed=6171778;  
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
RA Najarian R.C., Seeburg P.H., Wion K.L.;  
RT "The sequence of human serum albumin cDNA and its expression in E.  
RT coli.";  
RL Nucleic Acids Res. 9:6103-6114(1981).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczak A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RA Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RP SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.O.;  
RL (In) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RP SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Parrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RP SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RP SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RP VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]

RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,  
RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese alloalbumins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RP DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
human serum albumin whose precursor has an aberrant signal peptidase  
cleavage site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
Italy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RP VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
carboxyl-terminal variants of human serum albumin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
RN [24]  
RP VARIANT CASEBROOK ASN-518.  
RX MEDLINE=91316157; PubMed=1859851; DOI=10.1016/0925-4439(91)90023-3;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
albumin: albumin Casebrook (494 Asp-->Asn).";  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RX MEDLINE=92190239; PubMed=1347703; DOI=10.1016/0167-4838(92)90207-T;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two alloalbumins with identical electrophoretic mobility are produced  
Query Match 100.0%; Score 3103; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 5.1e-188;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHRFKDLGSENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHKSEVAHRFKDLGSENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 144  
QY 121 DVMCTAFHNEETFLKKLYIETARRHPYFVAPPELLFFAKRYKAAFTCCCAADKACLLP 180  
DB 145 DVMCTAFHNEETFLKKLYIETARRHPYFVAPPELLFFAKRYKAAFTCCCAADKACLLP 204

QY 181 KUDELURDEGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLT 240  
DB 205 KUDELURDEGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLT 264  
QY 241 VHTTECHGDLLECADDRLADLAKYICENQDSISSKLEKCEKCPLEKSHCIAEVENDEMPA 300  
DB 265 VHTTECHGDLLECADDRLADLAKYICENQDSISSKLEKCEKCPLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDYCKNYAEAKDVFGLMFLYEAARRHPDYSVVLRLRLAKTYETTTLEK 360  
DB 325 DLPSLAADFVESKDYCKNYAEAKDVFGLMFLYEAARRHPDYSVVLRLRLAKTYETTTLEK 384  
QY 361 CAADPHCEYAKVDFEKPPLVEEPQNLIKQNCLEFQOLGEYKFQNALVRYTKVPOVST 420  
DB 385 CAADPHCEYAKVDFEKPPLVEEPQNLIKQNCLEFQOLGEYKFQNALVRYTKVPOVST 444  
QY 421 PTLVSVSRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVSVSRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETPTFHADICTLSEKERQIKKQATLVELVKKHKPKAT 540  
DB 505 LVNRRPCFSALEVDETYVPKEFNAETPTFHADICTLSEKERQIKKQATLVELVKKHKPKAT 564  
QY 541 KEQLKAVMDDFAAFEVKECKKADDDKCTCFABEGKCLVAASQAALGL 585  
DB 565 KEQLKAVMDDFAAFEVKECKKADDDKCTCFABEGKCLVAASQAALGL 609  
RESULT 2  
Q68DN5 PRELIMINARY; PRT; 609 AA.  
AC Q68DN5;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein DKFZp779N1935.  
GN Name=DKFZp779N1935;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RG The German cDNA Consortium;  
RA Bloeker H., Boecker M., Brandt P., Mewes H.W., Weil B., Amid C.,  
RA Osanger A., Fobo G., Han M., Wiemann S.;  
RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.  
DR EMBL; CR749331; CAH18185.1; -;  
DR InterPro; IPR001703; AlphaFoldProt.  
DR InterPro; IPR000364; Serum albumin.  
DR Pfam; PF00273; Serum albumin; 3.  
DR PRINTS; PR00803; AFETOPROTEIN.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Hypothetical protein.  
SQ SEQUENCE 609 AA; 69402 MW; 3BA3AFF17BF99E94 CRC64;  
Query Match 99.2%; Score 3079; DB 2; Length 609;  
Best Local Similarity 99.1%; Pred. No. 1.7e-186;  
Matches 580; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHRFKDLGSENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHKSEVAHRFKDLGSENFKALVLIATAFYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 144

```
QY 121 DVMCTAFHDNBTFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180
Db 145 DVMCTAFHDNBTFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 204
QY 181 KLDELRLDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240
Db 205 KLDELRLDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPKLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVSKDVKCNKNAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVSKDVKCNKNAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEKPFLVEEPQNLIKQNCLEFQELGEYKFKQNALVRYTKKVPQVST 420
Db 385 CAADPHCEYAKVDFEKPFLVEEPQNLIKQNCLEFQELGEYKFKQNALVRYTKKVPQVST 444
QY 421 PTLVEVSNNLKGVSCKCKHPKAMPKCAEDYLSVLNQLCVLHKTTPVSDRVTKCCTES 480
Db 445 PTLVEVSNNLKGVSCKCKHPKAMPKCAEDYLSVLNQLCVLHKTTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVTPVKEFNFAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDVTPVKEFNFAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVNDDFAAFVEKCKCKADKTCFAEKGKLVAAASQAALGL 585
Db 565 KEQLKAVNDDFAAFVEKCKCKADKTCFAEKGKLVAAASQAALGL 609
```

## RESULT 3

```
ALBU MACMU
ID ALBU MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 05-JUL-2004 (Rel. 44, last annotation update)
DE Serum albumin precursor (Fragment).
GN Name=ALB;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
RT in bilirubin binding.";
RL proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR 103 EMBL; M90463; AAA36906.1; -.
DR PIR; A47391; A47391.
DR HSSP; P02768; 1B7B.
DR InterPro; IPR001703; Alphafetoprot.
DR Pfam; PF00264; Serum albumin.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Copper; lipid-binding; Metal-binding; Repeat; Signal.
FT NON TER 1
FT SIGNAL <1 10 By similarity.
FT PROPEP 11 16 By similarity.
FT CHAIN 17 600 Serum albumin.
FT DOMAIN 17 197 Albumin 1.
FT DOMAIN 204 389 Albumin 2.
FT DOMAIN 396 587 Albumin 3.
FT METAL 19 19 Copper (By similarity).
FT BINDING 256 256 Bilirubin (potential).
FT DISULFID 69 78 By similarity.
FT DISULFID 91 107 By similarity.
FT DISULFID 106 117 By similarity.
FT DISULFID 140 185 By similarity.
FT DISULFID 184 193 By similarity.
FT DISULFID 216 262 By similarity.
FT DISULFID 261 269 By similarity.
FT DISULFID 281 295 By similarity.
FT DISULFID 294 305 By similarity.
FT DISULFID 332 377 By similarity.
FT DISULFID 376 385 By similarity.
FT DISULFID 408 454 By similarity.
FT DISULFID 453 464 By similarity.
FT DISULFID 477 493 By similarity.
FT DISULFID 492 503 By similarity.
FT DISULFID 530 575 By similarity.
FT DISULFID 574 583 By similarity.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670B740B CRC64;
```

Query Match 94.8%; Score 2942; DB 1; Length 600;

Best Local Similarity 93.5%; Pred. No. 7.5e-178;

Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

```
QY 1 DAHSEVAHRFKDLGEENFKALVLIATAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 17 DTHSEVAHRFKDLGEENFKALVLIATAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 136
QY 121 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180
Db 137 DVMCTAFHDNEATFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 196
QY 181 KLDELRLDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 240
Db 197 KLDELRLDEGKASSAKQRLKCAASQKFGGERAFKAWAVARLSORFPKAEFAEYVKLVTDLTJK 256
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPKLLEKSHCIAEVENDEMPA 300
Db 257 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPKLLEKSHCIAEVENDEMPA 316
QY 301 DLPSLAADFVSKDVKCNKNAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 317 DLPSLAADFVSKDVKCNKNAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 376
QY 361 CAADPHCEYAKVDFEKPFLVEEPQNLIKQNCLEFQELGEYKFKQNALVRYTKKVPQVST 420
Db 377 CAADPHCEYAKVDFEKPFLVEEPQNLIKQNCLEFQELGEYKFKQNALVRYTKKVPQVST 436
QY 421 PTLVEVSNNLKGVSCKCKHPKAMPKCAEDYLSVLNQLCVLHKTTPVSDRVTKCCTES 480
```

Db 437 PTLVEVSRNLGVKAKCCCLPEAKRMPCAEDYLSVNLRLCVLHKEKTPVSEKVTKCCTES 496

QY 481 LVNRRPCFSALEVEDTYVPKEPNAETFTTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540

Db 497 LVNRRPCFSALEDAAYVPKAFNAETFTTFHADIMCTLSEKEKQVKTALVELVGHKPKAT 556

QY 541 KEOLKAVMDDFAAFEKCKCKADDKETCPFAEKGKLVAAQAAL 583

Db 557 KEOLKGVMDNFNAFVEKCKCKADDKETCPFAEKGKLVAAQAAL 599

RESULT 4

ALBU\_FELCA STANDARD; PRT; 608 AA.

AC P49064;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Serum albumin precursor (Allergen Fel d 2).

GN Name=ALB;

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI\_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;

RA Hilger C., Grigioni F., Kohnen M., Hentges F.;

RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";

RL Gene 169:295-296(1996).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- ALLERGEN: Causes an allergic reaction in human.

CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X84842; CAA59279.1; -.

DR PIR; J04660; S57632.

DR HSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; Serum albumin; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR SMART; PD002486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.

FT SIGNAL 1 18 By similarity.

FT PROPEP 19 24 By similarity.

FT CHAIN 25 608 Serum albumin.

FT DOMAIN 25 205 Albumin 1.

FT DOMAIN 212 397 Albumin 2.

FT DOMAIN 404 595 Albumin 3.

FT METAL 27 27 Copper.

FT DISULFID 77 86 By similarity.

FT DISULFID 99 115 By similarity.

FT DISULFID 114 125 By similarity.

FT DISULFID 148 193 By similarity.

FT DISULFID 192 201 By similarity.

FT DISULFID 224 270 By similarity.

FT DISULFID 269 277 By similarity.

FT DISULFID 289 303 By similarity.

FT DISULFID 302 313 By similarity.

FT DISULFID 340 385 By similarity.

FT DISULFID 384 393 By similarity.

FT DISULFID 416 462 By similarity.

FT DISULFID 461 472 By similarity.

FT DISULFID 485 501 By similarity.

FT DISULFID 500 511 By similarity.

FT DISULFID 538 583 By similarity.

FT DISULFID 582 591 By similarity.

SQ SEQUENCE 608 AA; 68659 MW; 07B629CACSF60E5F CRC64;

Query Match 84.4%; Score 2620; DB 1; Length 608;

Best Local Similarity 82.0%; Pred. No. 1.7e-157;

Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHKESEVAHRFKDLGSENFKALVIAFAQYLOQCQPFEDHVKLVNEVTEFAKTCVADESA 60

Db 25 EAHQSEIAHRFNDLGEHFRGLVAVFSQYLQCCPFEDHVKLVNEVTEFAKGVADQSA 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKQBPNERNECFLOHKDDNPRLVRPEV 120

Db 85 NCEKSLHELLGDKLCTVASLRDKYGEMADCCKEPERNECFLOHKDDNPFGVLTPEA 144

QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAAKAACL 180

Db 145 DAMCTAFHENEORFLGKYLIEIARRHPYFYAPPELLYAEYKGVFTECCAAADKAACLTP 204

QY 181 KLDELURDGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240

Db 205 KVDALREKVLASSAKERLKCASLOKFGGERAFKAWAVARLSQRPKAEFAEISKLVTDLAK 264

QY 241 VHTCECHGDLLECADRADLAKYICENODSISSKLKECEKPLLSKSHCIAEVENDMPA 300

Db 265 IHKECHGDLLECADRADLAKYICENODSISTKLKECCGKPVLEKSHCISEVERDELPA 324

QY 301 DLPSLAADPVESKDYCKNYAAKDYFLGMFLYVARRHPDYVSVLLLRLLAKTYETLEK 360

Db 325 DLPPLAADFVEDKEVKYQEAQVFLGTFLYVARRHPDYVSVLLLRLLAKTYETLEK 384

QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420

Db 385 CATDPPACYAHVDFEFKPLVEEPQNLKQNCLEFQELGEYGFQALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGVKAKCCCLPEAKRMPCAEDYLSVNLRLCVLHKEKTPVSEKVTKCCTES 480

Db 445 PTLVEVSRNLGVKAKCCCLPEAKRMPCAEDYLSVNLRLCVLHKEKTPVSEKVTKCCTES 504

QY 481 LVNRRPCFSALEVEDTYVPKEPNAETFTTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540

Db 505 LVNRRPCFSALEVEDTYVPKEPNAETFTTFHADICTLSEKERQIKKQTALVELVGHKPKAT 564

QY 541 KEOLKAVMDDFAAFEKCKCKADDKETCPFAEKGKLVAAQAAL 583

Db 565 KEOLKGVMDNFNAFVEKCKCKADDKETCPFAEKGKLVAAQAAL 607

RESULT 5

QYSG3 PRELIMINARY; PRT; 584 AA.

AC QYSG3

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Albumin (Fragment).

GN Name=alb;

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI\_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE=Liver; Swoboda I., Bohle B., Hauswirth A.W., Valent P., Reininger R.,

```
RA Rumpold H., Valenta R., Spitzauer S.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -.
DR HSSP; P02768; 1E7B.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 84.2%; Score 2614; DB 2; Length 584;
Best Local Similarity 81.6%; Pred. No. 4e-157;
Matches 476; Conservative 54; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 60
DB 1 EAHGSEIAHRNDLGESEHFRGLVAVFSQYLQCCPFEDHVKLVNVEVTEFANGCVADQSA 60

QY 61 NCDKSLHTLFGDKLCTVATLRETTGEMADCCAKQEPERNECFLOHKDNDPNLPRLVPEV 120
DB 1 NCKSLHFLFGDKLCTVATLRETTGEMADCCAKQEPERNECFLOHKDNDPNPGFQLVTEA 120

QY 121 DVMCTAFHDNETFLKYLVIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DAMCTAFHENEQRFLKYLVIARRHPYFYAPELLFYAEEYRGVFTCECAADKAACLTP 180

QY 181 KLDELRDGKASSAKORLKASLOKFGERAFKAVARLSORFPKABEAVSKLVTDLTK 240
DB 181 KYDALREKVLASSAKERUKCASLOKFGERAFKAVARLSORFPKADPAEISKLVTDLAK 240

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISKLKECCERPLEKSKHICIAEVENDEMPA 300
DB 241 IHKECCCHGDLLECADREDLAKYICENQDSISKLKECCGKPVLEKSHICSEVERDELPA 300

QY 301 DLPSLAADFVSKDVKCYAKADVFLGMFLYEVARRHPDYSVLLLLRLAKYITETLEKC 360
DB 301 DLPLAADFVEDKEYCKYQAEKADVFLGTLFLYESSRRHPDYSVLLLLRLAKYEATLEKC 360

QY 361 CAADPHCYAKVDFEKPPLVEEPONLJKONCELFEOQGEYKFNALLVRYTKVPQVST 420
DB 361 CATDDPPACYAHVDFEKPPLVEEPONLVKTNCELFKELGEYGFQVALLVRYTKVPQVST 420

QY 421 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLGKVGSKCKCHPEAERLSCAEDYLSVVLNRLCVLHKEKTPVSEKRVTKCCTES 480

QY 481 LVNRRPCFSALEVDYTPKGFNAETFTFHADICTLSEKERQIKQKOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYTPKGFSAETFTFHADLCTLPFAERQIKQKOTALVELLKHKPKAT 540

QY 541 KEQLKAVNMDFAAFVEKCKCKADDEKTCFAEKGKLVAAASQAAL 583
DB 541 EQQLKAVNMDFGSFVDKCAAEDEKTCFAEKGKLVAAASQAAL 583

RESULT 6
ALBU CANFA
ID _ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; O7705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN Name=ALB;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
```

```
FT DOMAIN 212 397 Alumin 2.
FT METAL 404 595 Copper (By similarity).
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
FT CONFLICT 1 26 MKVTFISLFFLFSSAYSRLVRREA -> MDT (in Ref. 2).
FT CONFLICT 146 146 A -> R (in Ref. 2).
FT CONFLICT 206 206 I -> T (in Ref. 2).
FT CONFLICT 349 349 V -> A (in Ref. 2).
FT CONFLICT 359 359 S -> A (in Ref. 1).
FT CONFLICT 448 448 V -> VV (in Ref. 1).
FT CONFLICT 474 474 D -> E (in Ref. 1).
SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;

Query Match 82.6%; Score 2562; DB 1; Length 608;
Best Local Similarity 79.8%; Pred. No. 8e-154;
Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 EAYKEIAHYNDLGEHFRGLVAFSYYLOQCPEFEDHVKLAKVETFEFAKCAAEBSGA 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNECFLOHKDDNPNLRLVRPEV 144

QY 121 DVMCTAFHNEBTFLLKYLIEIARRHPYFYAPPELLFPAKRYKAAFTCECCQAADKAACLPP 180
DB 145 DALCAAFQDNEQLFLGKYLIEIARRHPYFYAPPELLYAAQYKGVFAECCQAADKAACLGP 204

QY 181 KLDELDECKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPKAEPAEVSCLVTDLTJK 240
DB 205 KIEALREKVLLSAKERFKCASLQKFGGERAFKAWAVARLSQRPKADFAEISKVTDLTJK 264

QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCHIAEVENDEMPA 300
DB 265 VHKECCGDLLECADRADLAKYICENQDSISSKLKECCDKPVLEKSCQLAEVERDELPG 324

QY 301 DLPSLAADPVESKDYCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
DB 325 DLPSLAADPVESKDYCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLK 384

QY 361 CAAADPHCYAKVDFEPKPLVBEPPONLJKONCELPOLGEYKFONALLVRYTKVPQVST 420
DB 385 CATDDPTCYAKVDFEPKPLVBEPPONLJKONCELPOLGEYKFONALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVVLNOLCVLHHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVVLNOLCVLHHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCKADDKTCFAEEGKGLVAASQAAL 583
DB 565 DEQLKTVMGDFGAFVEKCKCKADDKTCFAEEGKGLVAASQAAL 607
```

## RESULT 7

```
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; -.
DR HSSP; P02768; IHKL.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; AlphaFetoprot.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
DR SEQUENCE 608 AA; 68225 MW; E5EAB828E1C66E54 CRC64;

Query Match 80.7%; Score 2504; DB 2; Length 608;
Best Local Similarity 76.3%; Pred. No. 3.7e-150;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHKEIAHRYNDLGEHFRGLVAFSYYLOQCPEFEDHVKLVNEVTEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCAIPTLRDSYGLADCCAKQBPNECFLOHKDDNPNLPPFVRPDA 144

QY 121 DVMCTAFHNEBTFLLKYLIEIARRHPYFYAPPELLFPAKRYKAAFTCECCQAADKAACLPP 180
DB 145 EAMCTSFQENAVTTFMGHVLHVEARRHPYFYAPPELLYAEKYSAIMTECCGEADKAACITP 204

QY 181 KLDELDECKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPKAEPAEVSCLVTDLTJK 240
DB 205 KLDALEKALASSVNRKLCSSLRQFGQRAFKAWAVARMSQKPPKADFAEITKLATDLTK 264

QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCPKLLEKSHCHIAEVENDEMPA 300
DB 265 LTBECCHGDLLECADRADLAKYICENQASISSKLQACCDKPVLSKSHCLSEVENDDLPA 324

QY 301 DLPSLAADPVESKDYCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLK 360
DB 325 DLPSLAADPVESKDYCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRLLAKTYETTLK 384

QY 361 CAAADPHCYAKVDFEPKPLVBEPPONLJKONCELPOLGEYKFONALLVRYTKVPQVST 420
DB 385 CAAEDPSACYGKVLDFEPKPLVBEPPONLJKONCELPOLGEYKFONALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVVLNOLCVLHHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVKSGCKKHPEAKRMPCAEDYLSVVLNOLCVLHHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 540
DB 505 VVERPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPKAT 564

QY 541 KEQLKAVMDDFAAFVEKCKCKADDKTCFAEEGKGLVAASQAAL 583
```









```

Db 204 KIETREKVLASSARQRLRCASIQKFGERALKAMSVARLSQKPPKAEFVEVTKLVTDLTk 263
Qy 241 VHTCECHGDLLECADDRLADLAKYICENQDSISSKKECECKPLLEKSHCHIAEVENDEMPA 300
Db 264 VHKECCHGDLLECADDRLADLAKYICDNQDTISSKKECCDKPLLEKSHCHIAEVENDAIPE 323
Qy 301 DLPSLAADFEVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETLEK 360
Db 324 NLPLTADFAEDKDVCKNYQEAADFLGSLFYESSRRHPYAVSVLLRLAKEYEATLEBC 383
Qy 361 CAADAPHECYAKVDFEKFPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPOVST 420
Db 384 CAKDDPHACYATVDFKLKHLVDEPNLIIKONCQDEKGEYGFQNALIVRYTKVPOVST 443
Qy 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 444 PTLVEVSRLSGKVGTRCTCKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTCKCTES 503
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHPKAT 540
Db 504 LVNRRPCFSALETPDETYVPKAFDEKLFTHADICTLPDTEKQIKKQTALVELLKHHPKAT 563
Qy 541 KEOLKAVMDDDFAAFVEKCKCADDKCTCFAEEGKCLVAASQAAL 583
Db 564 ESQKLTVMENFVAFVDDKCAADDKEACFAVEGPKLVVSTQTAL 606

RESULT 11
ALBU SHEEP
ID ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor.
GN Name=ALB;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]_TaxID=9940;
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90098888; PubMed=2602160;
RA Brown W.M., Dziaglewska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL Nucleic Acids Res. 17:10495-10495(1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17055; CAA34903.1; -.
CC PIR; S05936; A8SHS.
CC HSSP; P02768; 1HK1.
CC InterPro; IPR001703; Alphafetoprot.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; Serum albumin; 3.
CC PRINTS; P00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.

```

```

DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 78.4%; Score 2432.5; DB 1; Length 607;
Best Local Similarity 75.0%; Pred. No. 1.2e-145;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

Qy 1 DAHSEVAHRPKDGLGEENFKALVIAFAQYLQCCPFDEHVKLVNEVTEFAKTCVADESA 60
Db 25 DTHKSEIAHRFNDLGEENFQGLVLIASFQYLQCCPFDEHVKLVKELTEFAKTCVADESA 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCQEPERNECFLNHKDDSPDLPLK-KPEP 143
Qy 121 DVMCTAFHDNETFLKLYEIAERHPYFABELLFFAKRYKAAETECCEAADKACLLP 180
Db 144 DTLCAEFKADEKKFWGKYLYEVARRHPYFABELLYYANKYNGVFQECQABDKGACLLP 203
Qy 181 KLDELDEGKASSAKORLKCASLOKFGGERAFKAWAVARLSQRFKAEFAEVSGLVTDLTk 240
Db 204 KIDANREKVLASSARQRLRCASIQKFGERALKAMSVARLSQKFPKADFTDVTKIYTDLTk 263
Qy 241 VHTCECHGDLLECADDRLADLAKYICENQDSISSKKECECKPLLEKSHCHIAEVENDEMPA 300
Db 264 VHKECCHGDLLECADDRLADLAKYICDNQDTISSKKECCDKPLLEKSHCHIAEVDKDAVPE 323
Qy 301 DLPSLAADFEVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETLEK 360
Db 324 NLPLTADFAEDKDVCKNYQEAADFLGSLFYESSRRHPYAVSVLLRLAKEYEATLEDC 383
Qy 361 CAADAPHECYAKVDFEKFPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKVPOVST 420
Db 384 CAKDDPHACYATVDFKLKHLVDEPNLIIKONCELPOLGEYKFNALLVRYTKVPOVST 443
Qy 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 444 PTLVEVSRLSGKVGTRCTCKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTCKCTES 503
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHPKAT 540
Db 504 LVNRRPCFSDLTLDTEYVPKPFDEKFTFHADICTLPDTEKQIKKQTALVELLKHHPKAT 563
Qy 541 KEOLKAVMDDDFAAFVEKCKCADDKCTCFAEEGKCLVAASQAAL 583
Db 564 ESQKLTVMENFVAFVDDKCAADDKGGCFVLEGPKLIVASTQAAL 606

```

```

RESULT 12
ALBU RAT
ID ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN Name=Alb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223722; PubMed=7017712;
RA Sargent T.D., Yang M., Bonner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RN [2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin messenger
RT RNA.";
RL J. Biol. Chem. 252:6846-6855(1977).
RN [3]
RP SEQUENCE OF 25-222.
RX MEDLINE=78109429; PubMed=564345;
RA Isemura S., Ikenaka T.;
RT "Amino acid sequences of fragments I and II obtained by cyanogen
RT bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48(1978).
RN [4]
RP SEQUENCE OF 223-288 AND 572-608.
RX MEDLINE=76260153; PubMed=956149;
RA Isemura S., Ikenaka T.;
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
RT the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196(1976).
RN [5]
RP SEQUENCE OF 166-174.
RC TISSUE=Plasma;
RX MEDLINE=81194805; PubMed=2437111;
RA Caraway R.E., Mitra S.P., Cochran D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973(1987).
RN [6]
RP COPPER-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
RL Cancer Res. 38:3483-3486(1978).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and blood
CC flow (potential).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC

```

---

```

or send an email to license@isb-sib.ch.
-----
CC EMBL; V01222; CAA24532.1; -.
CC PIR; A93872; ABRTS.
CC HSSP; P02768; 1HK1.
CC Rat-heart-2DPAGE; P02770; -.
CC RGD; 2085; Alb.
CC InterPro; IPR001703; Alphafetoprot.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; Serum_albumin; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Copper; direct protein sequencing; Lipid-binding; Metal-binding;
CC Repeat; Signal.
CC SIGNAL 1 18
CC FT PROPEP 19 24
CC FT CHAIN 25 608
CC FT PEPTIDE 166 174
CC FT DOMAIN 25 205
CC FT DOMAIN 212 397
CC FT DOMAIN 404 595
CC FT METAL 27 27
CC FT DISULFID 77 86
CC FT DISULFID 99 115
CC FT DISULFID 114 125
CC FT DISULFID 148 193
CC FT DISULFID 192 201
CC FT DISULFID 224 270
CC FT DISULFID 269 277
CC FT DISULFID 289 303
CC FT DISULFID 302 313
CC FT DISULFID 340 385
CC FT DISULFID 384 393
CC FT DISULFID 416 462
CC FT DISULFID 461 472
CC FT DISULFID 485 501
CC FT DISULFID 500 511
CC FT DISULFID 538 583
CC FT DISULFID 582 591
CC FT VARIANT 262 262
CC FT CONFLICT 174 174 Y -> L (in Ref. 5).
CC SQ SEQUENCE 608 AA; 588497A282411AB7 CRC64;

Query Match 78.2%; Score 2426; DB 1; Length 608;
Best Local Similarity 73.4%; Pred. No. 3.2e-145;
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAKKSEVAHRPKDLGEENFKALVLIAPQYLQOCPPFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 EAHKSEIARHPKDLGEGHFKGLVLIAPSYLQKCPYBEHILKLVQEVDTFAKTCVADENAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPRLVREV 120
DB 85 NCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKOEPRNECFLOHKDDNPMLPPQFDEA 144
QY 121 DVMTCTADHNEETLKKYLVEIARRHYPFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 145 EAMCTSFQENPTSLGHLYHEVARHPFYFYAPPELLFYAEKYNEVLTOCCTESDKAACLT 204
QY 181 KLDELREDEGKASSAKQRLKCAQLQKGERAFKAWAVARLSQFFKAEFAEYKLVTLTK 240
DB 205 KLDVKEKALVAAVRQRMKSSMORFGERAFKAWAVARMSQFFNAEFAEITKLATDTVK 264
QY 241 VHTCCCHGDLLECDADRADLAKYTCENQDSISSKLKECCEKPLLEKSHICIAEVENDENPA 300
DB 265 INKECCHGDLLECDADRAELAKYMCENQATISSKLQACCDKPVQLQSKQLAEETHDNIPA 324
QY 301 DLPSIADDFVESKDVCNKYAEAKDVLGMLVEYARHPDYVSVLLRLRLAKTYETTLK 360
DB 325 DLPSIADDFVEDKEVCNKYAEAKDVLGTFLEYYSRRHPDYVSVLLRLRLAKKYEATLTK 384

```

QY 361 CAADPHCYAKVDFEKLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
Db 385 CAEGDPACVGTVLAEFQPLVEPKLVKTNCLYEKLGEYGFQNALVRYTKAPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEAARNLGRVGTCKCTLPEAQRLPVEDYLSALNRLCVLHEKTPVSEKVTCKCSGS 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHKKPAT 540  
Db 505 LVRRPCFSALTVDETYVPKEFAETFTPHSDICTLPEDEKQIKKOTALAELVKKHKKPAT 564  
QY 541 KEOLKAVMDPFAAFVEKCKKADDKETCFAEEGKLVAAQAAAL 583  
Db 565 EDQLKTVMGDFAQFVDKCKKAADKNCFCATEGPNLVARSKEAL 607  
RESULT 13  
ALBU\_PIG STANDARD; PRT; 605 AA.  
AC P08835; Q29018;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Serum albumin precursor (Fragment).  
GN Name=ALB;  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Liver;  
RC MEDLINE=89016582; PubMed=3174440;  
RA Baldwin G.S.; Weinstock J.;  
RT "Nucleotide sequence of porcine liver albumin.";  
RL Nucleic Acids Res. 16:9045-9045(1988).  
CC -! FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -! SUBCELLULAR LOCATION: Secreted.  
CC -! TISSUE SPECIFICITY: Plasma.  
CC -! SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -! SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X12422; CAA30970.1; -;  
CC EMBL; M36787; AAA30988.1; -;  
CC PIR; S01382; ABPGS.  
CC HSSP; P02768; 1E7H.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; Serum albumin.3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum\_albumin.1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
CC Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 16 By similarity.  
FT PROPEP 17 22 By similarity.  
FT CHAIN 23 605 Serum albumin.  
FT DOMAIN 23 202 Albumin 1.  
FT DOMAIN 209 394 Albumin 2.  
FT DOMAIN 401 592 Albumin 3.  
FT METAL 31 31 Copper (By similarity).

FT DISULFID 75 84 By similarity.  
FT DISULFID 97 113 By similarity.  
FT DISULFID 112 123 By similarity.  
FT DISULFID 145 190 By similarity.  
FT DISULFID 189 198 By similarity.  
FT DISULFID 221 267 By similarity.  
FT DISULFID 266 274 By similarity.  
FT DISULFID 286 300 By similarity.  
FT DISULFID 299 310 By similarity.  
FT DISULFID 337 382 By similarity.  
FT DISULFID 381 390 By similarity.  
FT DISULFID 413 459 By similarity.  
FT DISULFID 458 469 By similarity.  
FT DISULFID 482 498 By similarity.  
FT DISULFID 497 508 By similarity.  
FT DISULFID 535 580 By similarity.  
FT DISULFID 579 588 By similarity.  
FT CONFLICT 562 562 E -> D (in Ref. 1; AAA30988).  
SQ SEQUENCE 605 AA; 69410 MW; 3E55680DDIA1FAFF CRC64;  
Query Match 77.7%; Score 2411.5; DB 1; Length 605;  
Best Local Similarity 76.0%; Pred. No. 2.6e-144;  
Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;  
QY 1 DAHKEVAHRPKDILGEENFKALVLIAPQYLOOCPFEDHVKLVNEVTEFAKTCVADESAB 60  
Db 23 DTYKSEIAHRFKDLGELQVFKGLVLIAPSOHLQOCPYEEHVKLVREVTEFAKTCVADESAB 82  
QY 61 NCCKSLHTLFGDKLCTVATRLRETYGEMADCCAKQPERNECFLOHKDDNPNIPLRVPRPEV 120  
Db 83 NCCKSIHTLFGDKLCAIPSLREHYGDLADCCKEPERNECFLOHKDDNPDIPLK-KPDP 141  
QY 121 DVMCTAFHDNEETFLKLYEIAARRHPYPYAPPELLFFAKRYKAAFTCECCAAADKAAACLLP 180  
Db 142 VALCADFQDEQKFWGKLYEIAARRHPYPYAPPELLVYAIYKDFSECCQAADKAAACLLP 201  
QY 181 KLDELRDGKASSAKORLKASLOKFGRAFAMAVARLSORFPAEFAEVSRLVTDLTG 240  
Db 202 KIEHLREKVLTSAAKORLKASIQKFGRAFAMAVARLSORFPAEFAEVSRLVTDLTG 261  
QY 241 VHTCECHGDLLECADRADLAKYICENODSISSKLKECKECPLEKSHCIAEVENDEMPA 300  
Db 262 VHKCECHGDLLECADRADLAKYICENODTISTKLECKECPLEKSHCIAEVENDEMPA 321  
QY 301 DLPLSLAADPVESKDYCKNYAEAKDVFLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360  
Db 322 DLNPLEHDFVEDKEVKYKAEKDVFLGTFLVEYARRHPDYSVLLRLAKTYETTTLEDC 381  
QY 361 CAAADPHCYAKVDFEKLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
Db 382 CAKEDPPACVATVDFKQPLVDEPKNLIKONCELFQOLGEYKFNALLVRYTKVPOVST 441  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
Db 442 PTLVEAARNLGRVGTCKCTLPEAQRLPVEDYLSALNRLCVLHEKTPVSEKVTCKCSGS 501  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHKKPAT 540  
Db 502 LVNRRPCFSALTVDETYVPKEFAETFTPHADICTLPEDEKQIKKOTALVELVKKHKKPAT 561  
QY 541 KEOLKAVMDPFAAFVEKCKKADDKETCFAEEGKLV 576  
Db 562 EEQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKFV 597  
RESULT 14  
Q68NH7 PRELIMINARY; PRT; 607 AA.  
AC Q68NH7;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Albumin.



AC O35090; (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Serum albumin precursor.  
 GN Name=ALB;  
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
 OC Meriones.  
 OX NCBI\_TaxID=10047;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NGS IDR; TISSUE=Liver;  
 RX MEDLINE=98116663; PubMed=9455485;  
 RA Yoshida K., Seto-Ohshima A., Sinozara H.;  
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
 RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
 CC DNA Res. 4:351-354(1997).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/AFB/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AB006197; BAA21765.1; --  
 DR PIR; JC5838; JEW8.  
 DR HSP; P02768; 1E7B.  
 DR InterPro; IPR001703; Alphafetoprot.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
 FT SIGNAL 1 18 By similarity.  
 FT PROPEP 19 24 By similarity.  
 FT CHAIN 25 609 Serum albumin.  
 FT DOMAIN 25 206 Albumin 1.  
 FT DOMAIN 213 398 Albumin 2.  
 FT DOMAIN 405 596 Albumin 3.  
 FT METAL 28 28 Copper.  
 FT DISULFID 78 87 By similarity.  
 FT DISULFID 100 116 By similarity.  
 FT DISULFID 115 126 By similarity.  
 FT DISULFID 149 194 By similarity.  
 FT DISULFID 193 202 By similarity.  
 FT DISULFID 225 271 By similarity.  
 FT DISULFID 270 278 By similarity.  
 FT DISULFID 290 304 By similarity.  
 FT DISULFID 303 314 By similarity.  
 FT DISULFID 341 386 By similarity.  
 FT DISULFID 385 394 By similarity.  
 FT DISULFID 417 463 By similarity.  
 FT DISULFID 462 473 By similarity.  
 FT DISULFID 486 502 By similarity.  
 FT DISULFID 501 512 By similarity.  
 FT DISULFID 539 584 By similarity.  
 FT DISULFID 583 592 By similarity.  
 SEQUENCE 609 AA; 68940 MW; 9CASF97F67EF1A48 CRC64;

Query Match 76.9%; Score 2387; DB 1; Length 609;  
 Best Local Similarity 73.9%; Pred. No. 9.3e-143;  
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;  
 QY 2 AHKSEVAHFKDGLGNEFKALVLIFAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAN 61  
 DB 27 AHKSEIAHRYKDLGEKYFKGLVLYTFSQYLQKCSYEEHVHVKLVREVTDFASNCAKESAN 86  
 QY 62 CDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQEPNERNECFLOHKDDNNPRLVRPEVD 121  
 DB 87 CDKSLHTLFGDKLCSLPNFGKYEAMADCCAKQEPNERNECFLOHKDDNPQLPFPKRAEPD 146  
 QY 122 VMCTAFHDNEETFLKKLYLEYARRHPYFAPPELLFFAKRYKAAFTTECCOAAKACALLPK 181  
 DB 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGPPELLYLDKTYTAVLTCECAADKGACLTPK 206  
 QY 182 LDELDECKKASSAKORLKCSLQKFGERAFKAWAVARLSQRPFAEAEVSKLVLDLTKV 241  
 DB 207 LDALKEKALVSAVRQRLKCSSMKKFGERAFKAWAVARMSQTFPNADFAEITKLADLTKV 266  
 QY 242 HTECHGDLLECADRADLAKYICNQDSISSKLEKCEKPLLEKSHCIAEVENDEMPAD 301  
 DB 267 TOECHGDLLECADRAELAKYMCNQASISSKLQACCDKEMLKQSQCLAEVEHDDMPAD 326  
 QY 302 LPSLAADFVESKQCKNYAEAKQVPLGMFLYEVARRHPDYSVVLRLAKTYETTLKCC 361  
 DB 327 LPALTADFVEDKQVCKNYAEAKQVPLGMFLYEVARRHPDYSVVLRLAKTYETTLKCC 386  
 QY 362 AAADPHECYAKVDFDFKPLVEEPQNLIKONCELPQLGEYKFNQALLVRYTKKVPQVSTP 421  
 DB 387 AEADPHACYGHVDFDFKPLVEEPQNLVKSNCLEYKLGEGYFQNAVLYRYTKKAPQVSTP 446  
 QY 442 TLVEVSRNLGVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESL 481  
 DB 447 TLVEAARSILGRVGHCCALPEKKRLPCVEDYLSAILNRVCLLHEKTPVSEQVTKCCSGSL 506  
 QY 482 VNRRCPFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKOTALTVELVKKHPKATK 541  
 DB 507 VERRECFPSALPVDETYVPKEFNAETFTPHANICTLPEKEKQMEKQKOTALAEVKKHPQATE 566  
 QY 542 EQLKAVMDDFAAVFEKCKCKADDKCTCFABEGKKLVAASQAAAL 583  
 DB 567 EQLKAVMGDFAEFLKCKCKQEDKEACFTEGPKLVAESQKAL 608  
 RESULT 17  
 Q6B3Z0  
 ID Q6B3Z0 PRELIMINARY; PRT; 583 AA.  
 AC Q6B3Z0;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Serum albumin (Fragment).  
 GN Name=ESA;  
 OS Elephas maximus (Indian elephant).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.  
 OX NCBI\_TaxID=9783;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lazar J., Rasmussen B., Greenwood D.R., Bang I.-S., Prestwich G.D.;  
 RT "Elephant Albumin: A Multi-purpose Pheromone Shuttle.";  
 RL Chem. Biol. 0:0-0(2004).  
 DR EMBL; AY684122; AAT90502.1; --  
 DR GO; GO:0005615; C:extracellular space; IEA.  
 DR GO; GO:0005386; F:carrier activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.











**THIS PAGE BLANK (USPTO)**

Seq 18 Fused with Hgces4

- Fused

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:53 ; Search time 167 Seconds  
(without alignments)  
1662.838 Million cell updates/sec

Title: US-10-933-523-18

Perfect score: 3785

Sequence: 1 DAHKSEVAHRFKDLGEENFK.....leglreirsgsvqlcyvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 100 summaries

Database : A Geneseq 16Dec04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3156.5	83.4	763	8	ADL16712 Human stu
2	3153.4	83.3	838	7	ADL16494 Human alb
3	3153.4	83.3	838	7	ADH21273 Human alb
4	3152.8	83.3	876	7	ADL14940 Human alb
5	3152.8	83.3	876	7	ADH21285 Human alb
6	3152.3	83.3	819	7	ADL16517 Human alb
7	3152.3	83.3	819	7	ADH216519 Human alb
8	3152.3	83.3	819	7	ADH21806 Human alb
9	3152.3	83.3	819	7	ADH21808 Human alb
10	3152.2	83.3	809	7	ADL16560 Human alb
11	3152.2	83.3	794	7	ADL16451 Human alb
12	3152	83.3	774	7	ADL15070 Human alb
13	3152	83.3	774	7	ADL15068 Human alb
14	3151	83.2	774	7	ADL15006 Human alb
15	3151	83.2	774	7	ADL14959 Human alb
16	3151	83.2	774	7	ADL15004 Human alb
17	3151	83.2	774	7	ADL14949 Human alb
18	3151	83.2	774	7	ADL14960 Human alb
19	3151	83.2	868	7	ADL14923 Human alb
20	3151	83.2	868	7	ADH21275 Human alb
21	3151	83.2	874	7	ADL14922 Human alb
22	3151	83.2	874	7	ADH21274 Human alb
23	3151	83.2	880	7	ADL14925 Human alb
24	3151	83.2	880	7	ADL14983 Human alb
25	3151	83.2	880	7	ADL14978 Human alb

99 3133.2 82.8 1010 7 ADF16432 Adf16432 Human alb  
100 3133.2 82.8 1010 7 ADF16431 Adf16431 Human alb

ALIGNMENTS

RESULT 1  
ADL16712  
ID ADL16712 standard; protein; 763 AA.  
XX  
XX ADL16712;  
XX AC  
XX 17-JUN-2004 (first entry)  
XX  
DE Human stumulatory factor-interleukin 11 (IL-11) fusion protein.  
XX  
XX haemostatic; antianaemic; nephrotropic; cytostatic; anti-HIV;  
KW immunosuppressive; cell proliferation stimulatory factor;  
KW human serum albumin; HSA; CPSP; haematological disorder; hypochromia;  
KW hypochromic microcytic anaemia; anaemia; Platelet-less; HIV infection;  
KW cancer; renal failure; tissue transplant; organ transplant; human;  
KW interleukin 11; IL-11; fusion protein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN US2004063635-A1.  
XX  
XX 01-APR-2004.  
XX  
XX 26-JUN-2003; 2003US-00609346.  
XX  
XX 01-JUL-2002; 2002US-0392948P.  
XX  
PA (YUZZ/) YU Z.  
PA (FUYV/) FU Y.  
XX  
XX Yu Z, Fu Y;  
PI  
XX  
XX WPI; 2004-282521/26.  
DR N-PSDB; ADL16711.  
XX  
XX New isolated polynucleotide encoding a fusion protein formed between a  
PT human serum albumin (HSA) and a cell proliferation stimulatory factor  
PT (CPSP), useful for treating hematological disorders.  
XX  
XX Claim 34; SEQ ID NO 2; 65pp; English.  
PS  
XX The invention describes an isolated polynucleotide encoding a fusion  
CC protein formed between a human serum albumin (HSA) and a cell  
CC proliferation stimulatory factor (CPSP). Also described are: a  
CC recombinant vector comprising the isolated polynucleotide; a recombinant  
CC protein encoded by the polynucleotide; a recombinant cell comprising the  
CC recombinant vector; a composition comprising a combination of at least  
CC two different HSA/CPSP fusion proteins; a method for treating a patient  
CC with a CPSP, or haematological disorder; and a kit comprising a first  
CC fusion protein of HSA and a first CPSP, and a second fusion protein of  
CC HSA and a second CPSP. The polynucleotides, methods and compositions are  
CC useful for treating a patient needing CPSP or having haematological  
CC disorder, e.g. hypochromia, hypochromic microcytic anaemia and anaemia,  
CC platelet-less, HIV infection, cancer, renal failure and tissue/organ  
CC transplantation. This is the amino acid sequence of a fusion protein  
CC comprising human serum albumin (HSA) and cell stimulatory factor human  
CC interleukin 11 (IL-11).  
XX  
SQ Sequence 763 AA;  
Query Match 83.4%; Score 3156.5; DB 8; Length 763;  
Best Local Similarity 80.9%; Pred. No. 5e-73;  
Matches 619; Conservative 11; Mismatches 40; Indels 95; Gaps 9;  
XX  
XX 1 DAHSEVAHRFKDLGEENFKALVLI AFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

Db  
1 DAHSEVAHRFKDLGEENFKALVLI AFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60  
QY  
61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKKDDPNLRLVRPEV 120  
Db  
61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKKDDPNLRLVRPEV 120  
QY  
121 DVNCTAFHDNEETFLKKYLYEITARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180  
Db  
121 DVNCTAFHDNEETFLKKYLYEITARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180  
QY  
181 KLDELDEGKASSAKQRLKACSLQKFGERAFAKAWAVARLSQRFPKAFPAEVSKLVDLT 240  
Db  
181 KLDELDEGKASSAKQRLKACSLQKFGERAFAKAWAVARLSQRFPKAFPAEVSKLVDLT 240  
QY  
241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCKEPILLESKSHCIAEVENDEMPA 300  
Db  
241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCKEPILLESKSHCIAEVENDEMPA 300  
QY  
301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVVLRLAKTVETTLEK 360  
Db  
301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVVLRLAKTVETTLEK 360  
QY  
361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
Db  
361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420  
QY  
421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db  
421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY  
481 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
Db  
481 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540  
QY  
541 KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGLMSPRLVEPCSHALPQ 600  
Db  
541 KEOLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGLPGP-----PP 590  
QY  
601 G----LSP-----GOVIVRGLVLOEPKHFTVSLRQ----- 628  
Db  
591 GPPRVSPDRAELDSTVLLTRSL-LADTRQLAAQLRDKFPADGDHNLDSLFTLAMSAGAL 649  
QY  
629 -AAHAPVTLRASPAD-----RTLQWISRWGQKKLIS----- 658  
Db  
650 GALQLPGVLTRELRADLLSVLRHVQMLRRAGGSSLSKLTLEPELGTLOARLDLRLQLLMS 709  
QY  
659 -----APFLFYPO-----RPFVLLLFQEGGLKALN 685  
Db  
710 RLALPQPPPPAPPAPPAPPSSAWGGIRAAHAIL---GGHLTLTD 750  
RESULT 2  
ADFL6494  
ID ADF16494 standard; protein; 838 AA.  
XX  
XX ADF16494;  
AC ADF16494;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
XX Human albumin therapeutic fusion protein SeqID1591.  
XX  
KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX  
XX Chimeric.  
OS Homo sapiens.  
XX  
XX WO2003060071-A2.  
PN  
PD 24-JUL-2003.  
XX

PP 23-DEC-2002; 2002WO-US040891.  
 XX 21-DEC-2001; 2001US-0341811P.  
 PR 24-JAN-2002; 2002US-0350358P.  
 PR 28-JAN-2002; 2002US-0351360P.  
 PR 26-FEB-2002; 2002US-0359370P.  
 PR 28-FEB-2002; 2002US-0360000P.  
 PR 27-MAR-2002; 2002US-0367500P.  
 PR 08-APR-2002; 2002US-0370227P.  
 PR 10-MAY-2002; 2002US-0378950P.  
 PR 24-MAY-2002; 2002US-0382617P.  
 PR 28-MAY-2002; 2002US-0383123P.  
 PR 05-JUN-2002; 2002US-0385708P.  
 PR 10-JUL-2002; 2002US-0394625P.  
 PR 24-JUL-2002; 2002US-0398080P.  
 PR 09-AUG-2002; 2002US-0402131P.  
 PR 13-AUG-2002; 2002US-0402708P.  
 PR 18-SEP-2002; 2002US-0411355P.  
 PR 18-SEP-2002; 2002US-0411426P.  
 PR 02-OCT-2002; 2002US-0414984P.  
 PR 11-OCT-2002; 2002US-0417611P.  
 PR 23-OCT-2002; 2002US-0420246P.  
 PR 05-NOV-2002; 2002US-0423623P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.  
 PA (PRIN) PRINCIPIA PHARM CORP.  
 XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
 PI WPI; 2003-598517/56.  
 XX New albumin fusion protein, useful for preparing a composition for  
 PT treating diabetes mellitus.  
 XX Example 4; SEQ ID NO 1591; 24pp; English.  
 PS This invention relates to a novel albumin fusion protein having albumin  
 CC or biological activity. Human serum albumin is responsible for a  
 CC significant proportion of the osmotic pressure of serum and also  
 CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
 CC albumin to a therapeutic protein may increase shelf-life and stability of  
 CC the therapeutic protein. The albumin fusion protein of the invention may  
 CC allow production of compositions with antidiabetic activity whilst the  
 CC nucleotide sequence which encodes it may be useful for gene therapy. The  
 CC albumin fusion protein is useful for preparing a composition for treating  
 CC diabetes mellitus. The present sequence is the amino acid sequence of a  
 CC novel full-length human albumin therapeutic fusion protein of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
 XX Sequence 838 AA;  
 SQ  
 Query Match 83.3%; Score 3153.4; DB 7; Length 838;  
 Best Local Similarity 75.0%; Pred. No. 7e-73;  
 Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;  
 QY 1 DAHKSEVAHRFDKLGSENFKALVLAFAQYLOQCFEDHVKLVNVEVFAKTCVADESAE 60  
 Db 25 DAHKSEVAHRFDKLGSENFKALVLAFAQYLOQCFEDHVKLVNVEVFAKTCVADESAE 84  
 QY 61 NCDKSLHTLFGDKLTATRLTYGEMADCCAKQPERNECFLOHKDNPVLPLVRPEV 120  
 Db 85 NCDKSLHTLFGDKLTATRLTYGEMADCCAKQPERNECFLOHKDNPVLPLVRPEV 144  
 QY 121 DVWCTAFHNEETFLKKLYELARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db 145 DVWCTAFHNEETFLKKLYELARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
 QY 181 KLDELRLDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPFAEVSCLVDTLTK 240  
 Db 205 KLDELRLDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPFAEVSCLVDTLTK 264  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSSSKLKECCCKPLLEKSHCIAEVENDEMPA 300  
 Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSSSKLKECCCKPLLEKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRKAKTYETTLK 360  
 Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRKAKTYETTLK 384  
 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFKQVALLVRYTKVPQVST 420  
 Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFKQVALLVRYTKVPQVST 444  
 QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCABEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
 Db 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCABEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALVDEVTYVPKEFNAETFEHADICTLSEKEROIKKQATALVELVHKPKAT 540  
 Db 505 LVNRRPCFSALVDEVTYVPKEFNAETFEHADICTLSEKEROIKKQATALVELVHKPKAT 564  
 QY 541 KEQLKAVMDDFAAFEVEKCKCKADDKETCFABEGKKLVAASQAALGLM----- 586  
 Db 565 KEQLKAVMDDFAAFEVEKCKCKADDKETCFABEGKKLVAASQAALGLMDLLQLFLFVLL 624  
 QY 587 -----SPRLEV----- 592  
 Db 625 SGMGATGTLRTSLDPSLEIYKMKPEVKRREQLLAKNLALQNLNDIHOQYKILDVMLKGLFK 684  
 QY 593 -----PCSHALPOGLSPGOVLIIVRGLVLPQKHFTVSLRDOAHAPVTLRASPADR 643  
 Db 685 VLEDSRTVLTAAADVLPDGPFP-----QDEK-----LKDQAFSHV-VENTAFPGDV 727  
 QY 644 TLQ-----W--ISRWGQKKLISAPFLFYPQRFVLLLFQEGGLK--LALN 685  
 Db 728 VLRFPRIVHYFYFDHNSNNWLLIRWG---IS-----FCNQTVGVNQPHSPILSLM 774  
 QY 686 GQGIG-----ATSMNQAL-----EQLRELR 706  
 Db 775 AQELGISEKDSNFQNPFKIDRTFIPSTDPFKALREBEKRRKKEKKEIR 826  
 RESULT 3  
 ADH21793  
 ID ADH21793 standard; protein; 838 AA.  
 XX  
 AC ADH21793;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Human albumin/HCE1P80 fusion protein, SEQ ID NO:590.  
 XX  
 KW Fusion protein; human serum albumin; HSA; therapeutic protein;  
 KW shelf-life; in vitro biological activity; in vivo biological activity;  
 KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
 KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
 KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
 KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
 KW anorectic; ophthalmological; gene therapy.  
 XX  
 OS Chimeric.  
 OS Homo sapiens.  
 XX  
 PN WO2003059934-A2.  
 XX  
 PD 24-JUL-2003.  
 XX  
 PF 23-DEC-2002; 2002WO-US040892.  
 XX  
 XX 21-DEC-2001; 2001US-0341811P.  
 PR 24-JAN-2002; 2002US-0350358P.  
 PR 26-FEB-2002; 2002US-0359370P.  
 PR 28-FEB-2002; 2002US-0360000P.  
 PR 27-MAR-2002; 2002US-0367500P.  
 PR 08-APR-2002; 2002US-0370227P.  
 PR 10-MAY-2002; 2002US-0378950P.  
 PR 24-MAY-2002; 2002US-0382617P.  
 PR 28-MAY-2002; 2002US-0383123P.  
 PR 05-JUN-2002; 2002US-0385708P.  
 PR 10-JUL-2002; 2002US-0394625P.  
 PR 24-JUL-2002; 2002US-0398080P.  
 PR 09-AUG-2002; 2002US-0402131P.  
 PR 13-AUG-2002; 2002US-0402708P.  
 PR 18-SEP-2002; 2002US-0411355P.  
 PR 18-SEP-2002; 2002US-0411426P.  
 PR 02-OCT-2002; 2002US-0414984P.  
 PR 11-OCT-2002; 2002US-0417611P.  
 PR 23-OCT-2002; 2002US-0420246P.  
 PR 05-NOV-2002; 2002US-0423623P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.  
 PA (PRIN) PRINCIPIA PHARM CORP.  
 XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
 PI WPI; 2003-598517/56.  
 XX New albumin fusion protein, useful for preparing a composition for  
 PT treating diabetes mellitus.  
 XX Example 4; SEQ ID NO 1591; 24pp; English.  
 PS This invention relates to a novel albumin fusion protein having albumin  
 CC or biological activity. Human serum albumin is responsible for a  
 CC significant proportion of the osmotic pressure of serum and also  
 CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
 CC albumin to a therapeutic protein may increase shelf-life and stability of  
 CC the therapeutic protein. The albumin fusion protein of the invention may  
 CC allow production of compositions with antidiabetic activity whilst the  
 CC nucleotide sequence which encodes it may be useful for gene therapy. The  
 CC albumin fusion protein is useful for preparing a composition for treating  
 CC diabetes mellitus. The present sequence is the amino acid sequence of a  
 CC novel full-length human albumin therapeutic fusion protein of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
 XX Sequence 838 AA;  
 SQ  
 Query Match 83.3%; Score 3153.4; DB 7; Length 838;  
 Best Local Similarity 75.0%; Pred. No. 7e-73;  
 Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;  
 QY 1 DAHKSEVAHRFDKLGSENFKALVLAFAQYLOQCFEDHVKLVNVEVFAKTCVADESAE 60  
 Db 25 DAHKSEVAHRFDKLGSENFKALVLAFAQYLOQCFEDHVKLVNVEVFAKTCVADESAE 84  
 QY 61 NCDKSLHTLFGDKLTATRLTYGEMADCCAKQPERNECFLOHKDNPVLPLVRPEV 120  
 Db 85 NCDKSLHTLFGDKLTATRLTYGEMADCCAKQPERNECFLOHKDNPVLPLVRPEV 144  
 QY 121 DVWCTAFHNEETFLKKLYELARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db 145 DVWCTAFHNEETFLKKLYELARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
 QY 181 KLDELRLDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPFAEVSCLVDTLTK 240  
 Db 205 KLDELRLDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSQRPFAEVSCLVDTLTK 264

PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 03-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Haseltine WA;  
XX WPI; 2003-598501/56.  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX  
XX Disclosure; SEQ ID NO 590; 1086pp; English.  
XX  
XX The invention relates to fusion proteins comprising human serum albumin  
XX (ADH2I530) and a therapeutic polypeptide such as a therapeutic protein,  
XX antibody or peptide or their variants or fragments. The therapeutic  
XX protein may be fused to the N-terminus, the C-terminus or both termini of  
XX albumin via a linker. The albumin component of the fusion proteins  
XX prolongs the shelf-life and the in vitro and vivo biological activity of  
XX the proteins compared with those of the corresponding therapeutic  
XX proteins on their own. The invention also relates to nucleic acids  
XX encoding albumin fusion proteins, vectors and host cells comprising an  
XX albumin fusion protein nucleic acid, compositions and kits comprising an  
XX albumin fusion protein, the method of extending the shelf-life of a  
XX therapeutic protein by fusion with albumin, and the treatment of disease  
XX using an albumin fusion protein. The albumin fusion proteins may be used  
XX in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
XX related conditions. Specifically the albumin fusion proteins may be used  
XX to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
XX (especially neuropathy), retinopathy, cardiovascular disorders  
XX (especially heart disease, renal disorders and obesity). The proteins may  
XX also be used in a method of maintaining a basal glucose level in a  
XX patient and in a method for losing weight. The present sequence is  
XX related to the invention.  
XX  
XX Sequence 838 AA;  
XX  
XX Query Match 83.3%; Score 3153.4; DB 7; Length 838;  
XX Best Local Similarity 75.0%; Pred. No. 7e-73;  
XX Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;  
XX  
XX 1 DAHKSEVAHRFKOLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
XX 25 DAHKSEVAHRFKOLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
XX  
XX 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120  
XX 85 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 144  
XX  
XX 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
XX 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
XX  
XX 181 KLDELREGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPFPKAEFAEVSKLVTDLTK 240  
XX 205 KLDELREGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPFPKAEFAEVSKLVTDLTK 264  
XX  
XX 241 VHTCCCHGDLLECCADRADLAKYICENQDSTSSKLEKCECKPILKSHCIAEVENDEMPA 300  
XX 265 VHTCCCHGDLLECCADRADLAKYICENQDSTSSKLEKCECKPILKSHCIAEVENDEMPA 324  
XX  
XX 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360  
XX

Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTTLEKC 384  
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPEQLGEYKPFQNALVRYTKKVPQVST 420  
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELPEQLGEYKPFQNALVRYTKKVPQVST 444  
Qy 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEEDYLSVVNLQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEEDYLSVVNLQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYKPFNAETFTTHADICTLSEKERQIKKOTALTALVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYKPFNAETFTTHADICTLSEKERQIKKOTALTALVELVKKPKAT 564  
Qy 541 KEOLKAVMDDEFAFVEKCKCKADDDKETCFABEGKKLVAAASQAALGLM----- 586  
Db 565 KEOLKAVMDDEFAFVEKCKCKADDDKETCFABEGKKLVAAASQAALGLMDLLLOFLAFLFVLLL 624  
Qy 587 -----SPRLV----- 592  
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKKRREQLLAKNLAKLNDIHOQYKILDVMLKGLFK 684  
Qy 593 -----PCSHALPOGLSPGQVIVIRGLIVLOEPKHFTVSLRDQAAHAPVTLRASADR 643  
Db 685 VLEDSRTVLTAAADVLPDGPFP-----QDEK-----LKDAFASHV-VENTAFPGDV 727  
Qy 644 TLQ-----W--ISRWGOKKLISAPFLFPQPFVFLVLLFOEGGLK--LALN 685  
Db 728 VURFPRIYHYFDHNSNNLLIRWG-----IS-----FCNQTGVFNQGGHSPILSLM 774  
Qy 686 GQGLG-----ATSMNQAL-----BQLREL 706  
Db 775 AQELGISEKSNFQNFKIDRTFIPSTDPFOKALAEERKKRKEKKEIR 826  
XX  
XX RESULT 4  
XX ADF14940  
XX ID ADF14940 standard; protein; 876 AA.  
XX ADF14940;  
XX AC ADF14940;  
XX XX  
XX DT 12-FEB-2004 (first entry)  
XX XX  
XX DE Human albumin therapeutic fusion protein SeqID236.  
XX XX  
XX KW albumin fusion protein; albumin activity; human serum albumin;  
XX KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX KW gene therapy; diabetes mellitus; human.  
XX OS Chimeric.  
XX OS Homo sapiens.  
XX XX  
XX PN WO2003060071-A2.  
XX XX  
XX PD 24-JUL-2003.  
XX XX  
XX PF 23-DEC-2002; 2002WO-US040891.  
XX XX  
XX PR 21-DEC-2001; 2001US-0341811P.  
XX PR 24-JAN-2002; 2002US-0350358P.  
XX PR 28-JAN-2002; 2002US-0351360P.  
XX PR 26-FEB-2002; 2002US-0359370P.  
XX PR 28-FEB-2002; 2002US-0360000P.  
XX PR 27-MAR-2002; 2002US-0367500P.  
XX PR 08-APR-2002; 2002US-0370227P.  
XX PR 10-MAY-2002; 2002US-0378950P.  
XX PR 24-MAY-2002; 2002US-0382617P.  
XX PR 28-MAY-2002; 2002US-0383123P.  
XX PR 05-JUN-2002; 2002US-0385708P.  
XX PR 10-JUL-2002; 2002US-0394625P.  
XX PR 24-JUL-2002; 2002US-0398008P.  
XX PR 09-AUG-2002; 2002US-0402131P.  
XX PR 13-AUG-2002; 2002US-0402708P.



PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX  
XX WPI; 2003-598517/56.  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX  
XX  
XX Example 4; SEQ ID NO 236; 24pp; English.  
XX  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
XX Sequence 876 AA;

Query Match 83.3%; Score 3152.8; DB 7; Length 876;  
Best Local Similarity 70.4%; Pred. No. 7.9e-73;  
Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;

QY 1 DAHKEVAHRFDKLGEEFNKALVIAFAQYLOQCPFEDHVKLVNVTFAKTCVADESAG 60  
Db 25 DAHKEVAHRFDKLGEEFNKALVIAFAQYLOQCPFEDHVKLVNVTFAKTCVADESAG 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKYLIEARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLKYLIEARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264  
QY 241 VTECHGDLLECADDRADLAKYICENQDSISLKECECEKPLEKSHCIAEVENDEMPA 300  
Db 265 VTECHGDLLECADDRADLAKYICENQDSISLKECECEKPLEKSHCIAEVENDEMPA 324  
QY 301 DLPLSAAADPVESKDVCKNAEAKDVLGMLFLEYARRHPDYSVLLLLRLAKTYETTLK 360  
Db 325 DLPLSAAADPVESKDVCKNAEAKDVLGMLFLEYARRHPDYSVLLLLRLAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFDFKPLVEEPQNLIKONCELPQELGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFDFKPLVEEPQNLIKONCELPQELGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGVSKGCKCHGPEAKRMPCAEDVLSVVLNQLCVLHEKTPVSDRVTTCCTES 480  
Db 445 PTLVEVSRNLGVSKGCKCHGPEAKRMPCAEDVLSVVLNQLCVLHEKTPVSDRVTTCCTES 504

QY 481 LVNRRPCFSALEVDYTPKEFNAETTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYTPKEFNAETTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564  
QY 541 KEQLKAVMDDFAAFEKCKKADDKETCFABEGKKLVAASQAALGLMSPLREV----- 592  
Db 565 KEQLKAVMDDFAAFEKCKKADDKETCFABEGKKLVAASQAALGL-----LEVAETPTVPW 620  
QY 593 -----PCSHALPOGLSPGQVIVRGLVLQEPKHPT----- 622  
Db 621 RDAETGERLVCAQCPPTGTFVQRPCRDRSPITTCGP-----CPRHYTQFWMYLER 669  
QY 623 -----VSLRDQAANA----- 632  
Db 670 CRYCNVLGGEREEERACHATHNRACRGTGFFAHAGFCLEHASCPPGAGVIAGTPTSQN 729  
QY 633 -----PVTLRAS----- 639  
Db 730 TQCPQCPPTGTFSSSSSEQCPHRCNTALGLALNVPGSSSHDTLCTSGTGPLSTRVPG 789  
QY 640 -----FADRTLQWISR-----WG-----QKLLISAPFLFPYQ 666  
Db 790 ABECERAVIDFVAFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQLKL-----R 840  
QY 667 RPFVILLFQEGGLKLAINGOGLGATSNQOALEQLRELRSIG 709  
Db 841 RRLTELLGAQDQALLVRL-----LOALRVARMMPG 869

RESULT 5  
ADH21285  
ID ADH21285 standard; protein; 876 AA.  
XX AC ADH21285;  
XX DT 11-MAR-2004 (first entry)  
XX DE Human albumin/mutant TR6 fragment fusion protein, SEQ ID NO:82.  
XX KW Fusion protein; human serum albumin; HSA; therapeutic protein;  
KW shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
KW anorectic; ophthalmological; gene therapy.  
XX OS Synthetic.  
OS Chimeric.  
XX OS Homo sapiens.  
XX PN WO2003059934-A2.  
XX PD 24-JUL-2003.  
XX PF 23-DEC-2002; 2002WO-US040892.  
XX PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX

PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Haseltine WA;	
PI	WPI; 2003-598501/56.	
XX	New albumin fusion protein, useful for preparing a composition for	
PT	treating diabetes mellitus.	
PT	Disclosure; SEQ ID NO 82; 1086pp; English.	
XX	The invention relates to fusion proteins comprising human serum albumin	
CC	(ADH21530) and a therapeutic polypeptide such as a therapeutic protein,	
CC	antibody or peptide or their variants or fragments. The therapeutic	
CC	protein may be fused to the N-terminus, the C-terminus or both termini of	
CC	albumin via a linker. The albumin component of the fusion proteins	
CC	prolongs the shelf-life and the in vitro and vivo biological activity of	
CC	the proteins compared with those of the corresponding therapeutic	
CC	proteins on their own. The invention also relates to nucleic acids	
CC	encoding albumin fusion proteins, vectors and host cells comprising an	
CC	albumin fusion protein nucleic acid, compositions and kits comprising an	
CC	albumin fusion protein, the method of extending the shelf-life of a	
CC	therapeutic protein by fusion with albumin, and the treatment of disease	
CC	using an albumin fusion protein. The albumin fusion proteins may be used	
CC	in the treatment of metabolic/endocrine disorders, diabetes and diabetes-	
CC	related conditions. Specifically the albumin fusion proteins may be used	
CC	to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders	
CC	(especially neuropathy), retinopathy, cardiovascular disorders	
CC	(especially heart disease, renal disorders and obesity). The proteins may	
CC	also be used in a method of maintaining a basal glucose level in a	
CC	patient and in a method for losing weight. The present sequence is	
CC	related to the invention.	
XX	Sequence 876 AA;	
SQ	Query Match 83.3%; Score 3152.8; DB 7; Length 876;	
	Best Local Similarity 70.4%; Pred. No. 7.9e-73;	
	Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;	
Qy	1 DAHKSEVAHRPKDGENFKALVLIAPAYLIQQCFEDHVKLVNEVTEFAKTCVADSEAE 60	
Db	25 DAHKSEVAHRPKDGENFKALVLIAPAYLIQQCFEDHVKLVNEVTEFAKTCVADSEAE 84	
Qy	61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNNPLRLVRPEV 120	
Db	85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNNPLRLVRPEV 144	
Qy	121 DVMTCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180	
Db	145 DVMTCTAFHDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204	
Qy	181 KLDELREGKASSAKQRIKCSLQKFGERAPKAWAVARLSQRFPAKFAEVSKLVTDLTK 240	
Db	205 KLDELREGKASSAKQRIKCSLQKFGERAPKAWAVARLSQRFPAKFAEVSKLVTDLTK 264	
Qy	241 VHTCCCHDLLECCADRDALAKYICENODTSLSKLKCCCKPILKSHCTAEVNDSEMPA 300	
Db	265 VHTCCCHDLLECCADRDALAKYICENODTSLSKLKCCCKPILKSHCTAEVNDSEMPA 324	
Qy	301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTLK 360	
Db	325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTLK 384	
Qy	361 CAADPHCEYAKVDFEFPKLVVEEONLIKQNCLEFQELGEYKFNALLVRYTKVQPVST 420	
Db	385 CAADPHCEYAKVDFEFPKLVVEEONLIKQNCLEFQELGEYKFNALLVRYTKVQPVST 444	
Qy	421 PTLVEVSNLGVSKCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480	
Db	445 PTLVEVSNLGVSKCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 504	
Qy	481 LVNRRPCFSALEVDVETVYVPEKFNAEFTTFFHADICTLSEKERQIKKQTALVELVKHKPKAT 540	

Db	505 LVNRRPCFSALEVDVETVYVPEKFNAEFTTFFHADICTLSEKERQIKKQTALVELVKHKPKAT 564	
Qy	541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKLVAAQAALGLMSPLREV----- 592	
Db	565 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKLVAAQAALGL-----LEVAETPTYPW 620	
Qy	593 -----PCSHALPOGLSPGQVLIIVRGLVLOEPKHFT----- 622	
Db	621 RDAETGERLVCAQCPGPGTFVQRPQRDRDSTTCGP-----CPRRHYTFQWNYLER 669	
Qy	623 -----VSLRDOAAHA----- 632	
Db	670 CRYCNVLGGEREERARACHATHNRCACRTGFFAHAGFCLEHASCPPGAGVIAPGTPSQN 729	
Qy	633 -----PVTLRAS----- 639	
Db	730 TQCQCPGPGTFFSASSSSSSSQCPHRNCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPG 789	
Qy	640 -----FADRTLQWISR-----WG-----QKKLIASAPFLFYPO 666	
Db	790 ABECERAVIDFVAFQDISIKRLQLQLQALEAPEGNGTTPRAGRAALQLKL-----R 840	
Qy	667 RPEVLLVFOEGGLKALANGQGLGATSMNQALQLRELRLISG 709	
Db	841 RRLTELLGAGQDQALLVRL-----LQALRVARMPG 869	
RESULT 6		
ADF16517		
ID	ADF16517 standard; protein; 819 AA.	
XX	AC	ADF16517;
XX	DT	12-FEB-2004 (first entry)
XX	DE	Human albumin therapeutic fusion protein SeqID1614.
XX	KW	albumin fusion protein; albumin activity; human serum albumin;
XX	KW	serum osmotic pressure; shelf-life; stability; antidiabetic;
XX	KW	gene therapy; diabetes mellitus; human.
XX	OS	Chimeric.
XX	OS	Homo sapiens.
XX	PN	WO2003060071-A2.
XX	PD	24-JUL-2003.
XX	PF	23-DEC-2002; 2002WO-US040891.
XX	PR	21-DEC-2001; 2001US-0341811P.
PR	24-JAN-2002; 2002US-0350358P.	
PR	28-JAN-2002; 2002US-0351360P.	
PR	26-FEB-2002; 2002US-0359370P.	
PR	28-FEB-2002; 2002US-0360000P.	
PR	27-MAR-2002; 2002US-0367500P.	
PR	08-APR-2002; 2002US-0370227P.	
PR	10-MAY-2002; 2002US-0378950P.	
PR	24-MAY-2002; 2002US-0382617P.	
PR	28-MAY-2002; 2002US-0383123P.	
PR	05-JUN-2002; 2002US-0385708P.	
PR	14-JUL-2002; 2002US-0394625P.	
PR	24-JUL-2002; 2002US-0398008P.	
PR	09-AUG-2002; 2002US-0402131P.	
PR	13-AUG-2002; 2002US-0402708P.	
PR	18-SEP-2002; 2002US-0411355P.	
PR	18-SEP-2002; 2002US-0411426P.	
PR	02-OCT-2002; 2002US-0414984P.	
PR	11-OCT-2002; 2002US-0417611P.	
PR	23-OCT-2002; 2002US-0420246P.	
PR	05-NOV-2002; 2002US-042623P.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.

PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX (PRIN-) PRINCIPIA PHARM CORP.  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
DR New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX Example 4; SEQ ID NO 1614; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
XX Sequence 819 AA;  
Query Match 83.3%; Score 3152.3; DB 7; Length 819;  
Best Local Similarity 76.6%; Pred. NO. 7.2e-73;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;  
QY 1 DAHSEVAHRFDLGEENFKALVLAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFDLGEENFKALVLAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 144  
QY 121 DVMCTAFHNDNETFLKYLVEIARHPHYFAPELLFFAKRYKAAPTECCQAADKAACLLP 180  
DB 145 DVMCTAFHNDNETFLKYLVEIARHPHYFAPELLFFAKRYKAAPTECCQAADKAACLLP 204  
QY 181 KLDELRDGKASSAKORLKCASLOKFGGFAKAWARLSORFPAEAEVSKLVTDLTK 240  
DB 205 KLDELRDGKASSAKORLKCASLOKFGGFAKAWARLSORFPAEAEVSKLVTDLTK 264  
QY 241 VHTECHGDLLECCADRDALAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDENPA 300  
DB 265 VHTECHGDLLECCADRDALAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDENPA 324  
QY 301 DLPSLAADPVESKVKYAEAKDVFLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 360  
DB 325 DLPSLAADPVESKVKYAEAKDVFLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 384  
QY 361 CAADAPHECYAKVDFEFLVPEEPONLIKONCELFQEGYKFNALLVRYTKVPQVST 420  
DB 385 CAADAPHECYAKVDFEFLVPEEPONLIKONCELFQEGYKFNALLVRYTKVPQVST 444  
QY 421 PTLVSVNRLGKVGSKCKHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTTES 480  
DB 445 PTLVSVNRLGKVGSKCKHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTTES 504  
QY 481 LVNRRPCFSALVDETYVPKFNATFTPHADICTLSEKEROIKQTALVELVKKHKPAT 540  
DB 505 LVNRRPCFSALVDETYVPKFNATFTPHADICTLSEKEROIKQTALVELVKKHKPAT 564  
QY 541 KQOLKAVMDDDFAAFVEKCKADDKETCFABEGKLVAAASQAALGL-----MSPLREV 592  
DB 565 KQOLKAVMDDDFAAFVEKCKADDKETCFABEGKLVAAASQAALGLATGLTSLDPSLEI 624

QY 593 -----PCSHALPOGL 602  
DB 625 YKGMFEVRRREQLLAKNLQALNDIHQQYKILDMVLKGLFKVLEDSRTVLTAADVLPDGP 684  
QY 603 SPQGVIIIRGLVLQEPKHTVSLRDQAAAHAPVTLRASFAADRTLQ-----W- 647  
DB 685 FP-----QDEK-----LKDAFSHV-VENTAFFGDVLLRFFPRIVHYFDHNSNN 727  
QY 648 -ISRWGOKKLIISAPFLFPYQRFVLLFOEGGLK--LALNGQGLG----- 690  
DB 728 LLIRWG-----IS-----FCNQTGVFNQGHSPILSLMAQELGISSEKDSNFQNPFI 774  
QY 691 -----ATSMNQAL-----EQLRELR 706  
DB 775 DRTEFIPSTDFFQKALREERKKRKKESKRKEIR 807  
RESULT 7  
ID ADF16519 standard; protein; 819 AA.  
XX ADF16519;  
XX AC  
XX 12-FEB-2004 (first entry)  
XX Human albumin therapeutic fusion protein SeqID1616.  
XX albumin fusion protein; albumin activity; human serum albumin;  
XX serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.  
XX Chimeric.  
XX Homo sapiens.  
XX WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-0341811P.  
XX 24-JAN-2002; 2002US-0350358P.  
XX 26-FEB-2002; 2002US-0351360P.  
XX 26-FEB-2002; 2002US-0359370P.  
XX 27-MAR-2002; 2002US-0360000P.  
XX 08-APR-2002; 2002US-0367500P.  
XX 10-MAY-2002; 2002US-0370227P.  
XX 24-MAY-2002; 2002US-0382617P.  
XX 05-JUN-2002; 2002US-0383123P.  
XX 10-JUL-2002; 2002US-0394625P.  
XX 09-AUG-2002; 2002US-0402131P.  
XX 13-AUG-2002; 2002US-0402708P.  
XX 18-SEP-2002; 2002US-0411355P.  
XX 18-SEP-2002; 2002US-0411426P.  
XX 02-OCT-2002; 2002US-0414984P.  
XX 11-OCT-2002; 2002US-0417611P.  
XX 23-OCT-2002; 2002US-0420246P.  
XX 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.



CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity). The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.

XX SQ Sequence 819 AA;

Query Match 83.3%; Score 3152.3; DB 7; Length 819;  
Best Local Similarity 76.6%; Pred. No. 7.2e-73;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

Qy 1 DAHSEVAHREKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
Db 25 DAHSEVAHREKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDPNLRLVRPEV 144  
Qy 121 DVMCTAFHDNEETFLKKYLYBIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYLYBIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
Qy 181 KLDELDRGKASSAKQRIKCSAQKFGGERAFKAWAVARLSORFFKAEFAEYKLVTLDTLK 240  
Db 205 KLDELDRGKASSAKQRIKCSAQKFGGERAFKAWAVARLSORFFKAEFAEYKLVTLDTLK 264  
Qy 241 VHTCCGDLLECCADDRADLAKYICENODSSSKLKECCPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGDLLECCADDRADLAKYICENODSSSKLKECCPELLEKSHCIAEVENDEMPA 324  
Qy 301 DPLSLAADFVESKDVCKNYAEAKVFLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 360  
Db 325 DPLSLAADFVESKDVCKNYAEAKVFLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 384  
Qy 361 CAADDPHECYAKVDFEPFLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420  
Db 385 CAADDPHECYAKVDFEPFLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 444  
Qy 421 PTLVEVSRLGKVGSKCKHPKAKMPCAEEDVLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKHPKAKMPCAEEDVLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKRIQKKQTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKRIQKKQTALVELVKHKPKAT 564  
Qy 541 KEQLKAVNDDFAAFVEKCKADDKTCFAEBGKKLVAASQAALGI-----MSPLREV 592  
Db 565 KEQLKAVNDDFAAFVEKCKADDKTCFAEBGKKLVAASQAALGIATGLTRTSLDPSLEI 624  
Qy 593 -----QDEK-----LKDASHV-VENTAFFGDVVRFPRIYVHYFDHNSWN 727  
Db 625 YKMFVEVRREQLLAKNLALQNDTHQYKILDVLMKGLFKVLEDSRVTILTAADVLPDGP 684  
Qy 603 SPQGVIIIVRGLVQLPQKFTVSLRQAAHAPVTLRASFADRTLQ-----W- 647  
Db 685 FP-----QDEK-----LKDASHV-VENTAFFGDVVRFPRIYVHYFDHNSWN 727  
Qy 648 -ISRNGQKLLISAPFLFPQRFVEVLLLFQEBGLK--LALNGQGLG----- 690  
Db 728 LLIRNG-----IS-----FCNOTGVFNQGHSPILSLMAQELGISEKDSNFQPFKI 774  
Qy 691 -----ATSMNQAL-----EQLRELR 706  
Db 775 DRTEFIPSTDPPQKALREBEKKRKKKEIR 807

RESULT 9  
ADH21808

ID ADH21808 standard; protein; 819 AA.

AC ADH21808;

DT 11-MAR-2004 (first entry)

DE Human albumin/HCE1P80 fragment fusion protein, SEQ ID NO:605.

XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
XX shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiatic;  
anorectic; ophthalmological; gene therapy.

OS Chimeric.

OS Homo sapiens.

WO2003059934-A2.

PD 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US040892.

XX 21-DEC-2001; 2001US-0341811P.

PR 24-JAN-2002; 2002US-0350358P.

PR 26-FEB-2002; 2002US-0359370P.

PR 28-FEB-2002; 2002US-0360000P.

PR 27-MAR-2002; 2002US-0367500P.

PR 08-APR-2002; 2002US-0370227P.

PR 10-MAY-2002; 2002US-0378950P.

PR 24-JUL-2002; 2002US-0398008P.

PR 09-AUG-2002; 2002US-0402131P.

PR 13-AUG-2002; 2002US-0402708P.

PR 18-SEP-2002; 2002US-0411355P.

PR 02-OCT-2002; 2002US-0414984P.

PR 11-OCT-2002; 2002US-0417611P.

PR 23-OCT-2002; 2002US-0420246P.

PR 05-NOV-2002; 2002US-0423623P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Haseltine WA;

XX WPI; 2003-598501/56.

XX New albumin fusion protein, useful for preparing a composition for

PT treating diabetes mellitus.

XX Disclosure; SEQ ID NO 605; 1086pp; English.

XX The invention relates to fusion proteins comprising human serum albumin  
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments. The therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini  
CC of albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein by fusion with albumin, and the treatment of a  
CC disease using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity). The proteins may  
CC also be used in a method of maintaining a basal glucose level in a

CC	patient and in a method for losing weight. The present sequence is	
CC	related to the invention.	
XX		
XX	Sequence 819 AA;	
XX	Query Match 83.3%; Score 3152.3; DB 7; Length 819;	
XX	Best Local Similarity 76.6%; Pred. No. 7.2e-73;	
XX	Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;	
Qy	1 DAHSEVAHFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES	60
Db	25 DAHSEVAHFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES	84
Qy	61 NCDKSLHTLFGDKLCTVATIRYTGEMADCAKQEPERNECFLOHODDNNLRLVRPEV	120
Db	85 NCDKSLHTLFGDKLCTVATIRYTGEMADCAKQEPERNECFLOHODDNNLRLVRPEV	144
Qy	121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCOAAKAAACLLP	180
Db	145 DVMTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCOAAKAAACLLP	204
Qy	181 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK	240
Db	205 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK	264
Qy	241 VHTCCGHDLLLECADRADLAKYICENQDISSKLKECCCKPILLEKSHCIAEVENDEMPA	300
Db	265 VHTCCGHDLLLECADRADLAKYICENQDISSKLKECCCKPILLEKSHCIAEVENDEMPA	324
Qy	301 DLPSLAADFVESKDVCKNYAEAKDVLGMEFYEARHPDYSVVLLRLAKTYETTLK	360
Db	325 DLPSLAADFVESKDVCKNYAEAKDVLGMEFYEARHPDYSVVLLRLAKTYETTLK	384
Qy	361 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFQGLGEYKFNQALLVRYTKVPQVST	420
Db	385 CAADPHCEYAKVDFEKLVEEPQNLIKONCELFQGLGEYKFNQALLVRYTKVPQVST	444
Qy	421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES	480
Db	445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES	504
Qy	481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVKHKPKAT	540
Db	505 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVKHKPKAT	564
Qy	541 KEQLKAYMDDFAAFVEKCCCKADKCTCFABEGKCLVAASQAALGL-----MSRPLEV	592
Db	565 KEQLKAYMDDFAAFVEKCCCKADKCTCFABEGKCLVAASQAALGLATGLTSLDPSLEI	624
Qy	593 -----PCSHALPQGL 602	
Db	625 YKMFVEVRREQLLAKNLQALNDIHOQYKILDVLMKGLFKVLBDSRTVLTAADVLDGP	684
Qy	603 SPQGVIIVRGLVQEPKHTVSLRDQAHAHVTLRASFAEDTLQ-----W- 647	
Db	685 FP-----ODEK-----LKDAFSHV-VENTAFGDVVLPRPIVHYVDHNSWN	727
Qy	648 -ISWGQKLTISAPFLYPPQRFVFLVLLFOEGGLK--LALNGQGLG-----690	
Db	728 LLIRWG---IS-----FCNQTVFNQGHSPILSLMAQELGLISEKDSNPFQNFKI	774
Qy	691 -----ATSMNQAL-----BQLREL 706	
Db	775 DRTEFIPSTPPQKALREEEKRRKKEIR 807	
RESULT 10		
ADFL6560		
ID	ADFL6560 standard; protein; 794 AA.	
XX		
AC	ADFL6560;	
XX		
DT	12-FEB-2004 (first entry)	

XX	Human albumin therapeutic fusion protein SeqID1657.	
DE	albumin fusion protein; albumin activity; human serum albumin;	
XX	serum osmotic pressure; shelf-life; stability; antidiabetic;	
KW	gene therapy; diabetes mellitus; human.	
XX	Chimeric.	
OS	Homo sapiens.	
XX	WO2003060071-A2.	
PN		
XX	24-JUL-2003.	
PD		
XX	23-DEC-2002; 2002WO-US040891.	
PF		
XX	21-DEC-2001; 2001US-0341811P.	
PR	24-JAN-2002; 2002US-0350358P.	
PR	28-JAN-2002; 2002US-0351360P.	
PR	26-FEB-2002; 2002US-0359370P.	
PR	28-FEB-2002; 2002US-0360000P.	
PR	27-MAR-2002; 2002US-0367500P.	
PR	08-APR-2002; 2002US-0370227P.	
PR	10-MAY-2002; 2002US-0378950P.	
PR	24-MAY-2002; 2002US-0382617P.	
PR	28-MAY-2002; 2002US-0383123P.	
PR	05-JUN-2002; 2002US-0385708P.	
PR	10-JUL-2002; 2002US-0394625P.	
PR	24-JUL-2002; 2002US-0398008P.	
PR	09-AUG-2002; 2002US-0402131P.	
PR	13-AUG-2002; 2002US-0402708P.	
PR	18-SEP-2002; 2002US-0411355P.	
PR	18-SEP-2002; 2002US-0411426P.	
PR	02-OCT-2002; 2002US-0414984P.	
PR	11-OCT-2002; 2002US-0417611P.	
PR	23-OCT-2002; 2002US-0420246P.	
PR	05-NOV-2002; 2002US-0423623P.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	(DELZ ) DELTA BIOTECHNOLOGY LTD.	
PA	(PRIN-) PRINCIPIA PHARM CORP.	
XX	Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;	
XX	WPI; 2003-598517/56.	
DR	New albumin fusion protein, useful for preparing a composition for	
XX	treating diabetes mellitus.	
XX	Example 4; SEQ ID NO 1657; 24pp; English.	
PS	This invention relates to a novel albumin fusion protein having albumin	
CC	or biological activity. Human serum albumin is responsible for a	
CC	significant proportion of the osmotic pressure of serum and also	
CC	functions as a carrier of endogenous and exogenous ligands. The fusion of	
CC	albumin to a therapeutic protein may increase shelf-life and stability of	
CC	the therapeutic protein. The albumin fusion protein of the invention may	
CC	allow production of compositions with antidiabetic activity whilst the	
CC	nucleotide sequence which encodes it may be useful for gene therapy. The	
CC	albumin fusion protein is useful for preparing a composition for treating	
CC	diabetes mellitus. The present sequence is the amino acid sequence of a	
CC	novel full-length human albumin therapeutic fusion protein of the	
CC	invention. Note: The sequence data for this patent did not form part of	
CC	the printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/publishedpct_sequences	
XX		
SQ	Sequence 794 AA;	
Query Match	83.3%; Score 3152.2; DB 7; Length 794;	
Best Local Similarity	76.3%; Pred. No. 6.9e-73;	
Matches	617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;	
Qy	1 DAHSEVAHFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES	60

Db 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAPKAWAVARLSQRPFKABFAEVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAPKAWAVARLSQRPFKABFAEVSKLVTDLTK 264  
Qy 241 VHTTECHGDLLECCADDDRADLAKYICENQDSTSSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Db 265 VHTTECHGDLLECCADDDRADLAKYICENQDSTSSKLKECCPKLLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 384  
Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKCHPEAKMPCAEADYLSVNLNOLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKMPCAEADYLSVNLNOLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVTPKFEFNAETFTFHADICTLSEKERQIKKQTLALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDVTPKFEFNAETFTFHADICTLSEKERQIKKQTLALVELVKHKPKAT 564  
Qy 541 KEQLKAVMDPFAFVEKCKCKADDETCFAEKGKLVAAASQAALGLM-----SPRLEV 592  
Db 565 KEQLKAVMDPFAFVEKCKCKADDETCFAEKGKLVAAASQAALGLMAFTEHSPLTPHRRD 624  
Qy 593 PCSHAL-----PQGLSPG----- 605  
Db 625 LCSRSIWLARKIRSDLTALTESYVKGHGLNKNINLDSADGMPVASTDQWSELTEABRLQE 684  
Qy 606 -----QVLIIVGLVLQEPKHFVSLRD--QAAHAPVTLRASFA----- 641  
Db 685 NLQAYRTHVLUAR--LLEDQVHFPTPTGDFHQAIHTLLQVAFAPVQIEBMLMILEYKI 743  
Qy 642 -----DRTLQWISRWGKKLISAPFLFYPPORFFVLLLFQEGGLKIA 683  
Db 744 PRNEADGMPINVGDLPEKKL-----WGLKVL-----QE----- 773  
Qy 684 LMGQGLGATSNQQAQLRELRLISGSVQ 712  
Db 774 -----LSQWTVRSIHLDRFTSSHQ 792

RESULT 11

ADFL6451  
ID ADF16451 standard; protein; 809 AA.

XX AC ADF16451;

XX DT 12-FEB-2004 (first entry)

XX DE Human albumin therapeutic fusion protein SeqID1548.

XX KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.

XX OS Chimeric.  
XX OS Homo sapiens.

PN WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 29-AUG-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
treating diabetes mellitus.  
XX Example 4; SEQ ID NO 1548; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
or biological activity. Human serum albumin is responsible for a  
significant proportion of the osmotic pressure of serum and also  
functions as a carrier of endogenous and exogenous ligands. The fusion of  
albumin to a therapeutic protein may increase shelf-life and stability of  
the therapeutic protein. The albumin fusion protein of the invention may  
allow production of compositions with antidiabetic activity whilst the  
nucleotide sequence which encodes it may be useful for gene therapy. The  
albumin fusion protein is useful for preparing a composition for treating  
diabetes mellitus. The present sequence is the amino acid sequence of a  
novel full-length human albumin therapeutic fusion protein of the  
invention. Note: The sequence data for this patent did not form part of  
the printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

XX Sequence 809 AA;

Query Match 83.3%; Score 3152.2; DB 7; Length 809;

Best Local Similarity 76.3%; Pred. No. 7.1e-73;  
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204

QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264  
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 300  
Db 265 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLRLRAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLRLRAKTYETTLK 384  
QY 361 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420  
Db 385 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 444  
QY 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDETVYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 564  
QY 541 KEOLKAVMDDFAAVFVEKCKKADDKETCFABEKGKLVAAQOALGLM-----SPRLEV 592  
Db 565 KEOLKAVMDDFAAVFVEKCKKADDKETCFABEKGKLVAAQOALGLMAFTEHSPLTPHRRD 624  
QY 593 PCSHAL-----PQGLSPG----- 605  
Db 625 LCSRSIWLARKISDLTALTESVYVHQGLNKNINLDSADGMPVASTDQWSELTEARLQE 684  
QY 606 -----QVIIVRGLVLOEPKHFTVSLRD--QAAHAPVTLRASFA----- 641  
Db 685 NLOAVRTFHVLLAR-LLEDQOVHFTTEGDFHOAIHTLLIQVAAFAYQIEELMILLEYKI 743  
QY 642 -----DRTLQWISRWGOKKLISAPFLFYPQRPFEVLLLFQEGGLKLA 683  
Db 744 PRNEADGMPINVGGLGFEKKL-----WGLKVL-----QE----- 773  
QY 684 LNQGLGATSMNQALEQKRELKISGSVQ 712  
Db 774 -----LSQWTVRSIHDLAFISSHQ 792

RESULT 12

ADFI5070  
ID ADFI5070 standard; protein; 774 AA.  
XX AC ADFI5070;  
XX DT 12-FEB-2004 (first entry)  
XX DE Human albumin therapeutic fusion protein SeqID366.  
XX KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX OS Chimeric.  
XX OS Homo sapiens.  
XX PN WO2003060071-A2.  
XX PD 24-JUL-2003.  
XX PF 23-DEC-2002; 2002WO-US040891.  
XX PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
PI WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX Example 4; SEQ ID NO 366; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX SQ Sequence 774 AA;  
Query Match 83.3%; Score 3152; DB 7; Length 774;  
Best Local Similarity 80.3%; Pred. No. 6.7e-73;  
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;  
QY 1 DAHKSVAHRFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADES 60  
Db 25 DAHKSVAHRFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADES 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYVAPELLFFAKRYKAAFTCCQAADKAACLLP 180  
Db 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYVAPELLFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264  
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 300  
Db 265 VHTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLRLRAKTYETTLK 360



Db 325 DLPSLAADFVSKDVKCNKAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 384  
Qy 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETYPKFNABTFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETYPKFNABTFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 564  
Qy 541 KEQLKAVMDDFAAVFEKCKKADDDKTCFAEBGKGLVAASQAALGLMSPRL----- 590  
Db 565 KEQLKAVMDDFAAVFEKCKKADDDKTCFAEBGKGLVAASQAALGLMSPRL----- 624  
Qy 591 -----EVPKSHA-----LPQGLS 603  
Db 625 LLEAKAEQITTCGAHCSLNEQITVPTDKVNFYAKKEMEVGQQAQAVEVWQGLALLSEAVL 684  
Qy 604 PGQVIVRGLVQLQP-----KHFTVSLRDQAAHAPVTLRASFAVRTLQWISRW 651  
Db 685 RGQALLVQSSQWEPQLHVDKAVSGLSLATLLR-----ALRA----- 723  
Qy 652 GQKLIASAP-----PLFPYQFFEVLLLFQSGGLKL 682  
Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFLRGKUL 764

RESULT 13  
ADFL5068  
ID ADFL5068 standard; protein; 774 AA.  
XX AC ADFL5068;  
DT 12-FEB-2004 (first entry)  
XX DE Human albumin therapeutic fusion protein SeqID364.  
XX DE albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX Chimeric.  
OS Homo sapiens.  
XX WO2003060071-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040891.  
XX 21-DEC-2001; 2001US-0341811P.  
XX 24-JAN-2002; 2002US-0350358P.  
XX 28-JAN-2002; 2002US-0351360P.  
XX 26-FEB-2002; 2002US-0359370P.  
XX 28-FEB-2002; 2002US-0360000P.  
XX 27-MAR-2002; 2002US-0367500P.  
XX 08-APR-2002; 2002US-0370227P.  
XX 10-MAY-2002; 2002US-0378950P.  
XX 24-MAY-2002; 2002US-0382617P.  
XX 28-MAY-2002; 2002US-0383123P.  
XX 05-JUN-2002; 2002US-0385708P.  
XX 10-JUL-2002; 2002US-0394625P.  
XX 24-JUL-2002; 2002US-0398008P.  
XX 09-AUG-2002; 2002US-0402131P.  
XX 13-AUG-2002; 2002US-0402708P.  
XX 18-SEP-2002; 2002US-0411355P.  
XX 18-SEP-2002; 2002US-0411426P.  
XX 02-OCT-2002; 2002US-0414984P.

PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PT Example 4; SEQ ID NO 364; 24pp; English.  
PS This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX SQ Sequence 774 AA;  
Query Match 83.3%; Score 3152; DB 7; Length 774;  
Best Local Similarity 80.3%; Pred. No. 6.7e-73;  
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;  
Qy 1 DAHSEVAHREFKOLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHREFKOLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRREV 144  
Qy 121 DVMCTAFHDNEETFLKKYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 204  
Qy 181 KLDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFKAEFAEVSCLVTDLTG 240  
Db 205 KLDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFKAEFAEVSCLVTDLTG 264  
Qy 241 VHTCCGHDLLLECADDDRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGHDLLLECADDDRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVSKDVKCNKAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVSKDVKCNKAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 384  
Qy 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETYPKFNABTFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETYPKFNABTFTFHADICTLSEKEROIKKQOTALVELVGHKPKAT 564

```
QY 541 KEOLKAVMDDDFAAFVEKCKKADDETCFAEEGKKLVAAQAALGLMSPRLL----- 590
Db 565 KEOLKAVMDDDFAAFVEKCKKADDETCFAEEGKKLVAAQAALGLAPPRLICDSRVLYRY 624
QY 591 -----EVPSCSHA-----LPQGLS 603
Db 625 LLEAKEAEQITTCGAHCSSLNEQITVPDTKVNFWYKMEVEGQQAQVEMVQGLLSEAVL 684
QY 604 PGQVIVIRGLVLOEP-----KHTVSLRDOQAHAHVTLRASPADRTLQWISRW 651
Db 685 RGQALLVQSSQPWEPLQLQHVHDKAVGLRSLTLLR-----ALRA----- 723
QY 652 GQKKLISAP-----FLFYPRFFEVLLLFQEGGKL 682
Db 724 -QKEAISPPDAASAPLRTITADTF-----RKLPFRVSNFURGKLKL 764

RESULT 14
ADF15006
ID ADF15006 standard; protein; 774 AA.
AC ADF15006;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human albumin therapeutic fusion protein SeqID302.
XX
KW albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003060071-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040891.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 28-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
XX
DR WPI; 2003-598517/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
```

```
XX
PS
XX
CC This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is the amino acid sequence of a
CC novel full-length human albumin therapeutic fusion protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/publishedpct_sequences
XX
SQ Sequence 774 AA;
Query Match 83.2%; Score 3151; DB 7; Length 774;
Best Local Similarity 80.3%; Pred. No. 7.1e-73;
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;
QY 1 DAHKSEVAHRFKDLGGENFKALVLIQVLAQVLOQCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHKSEVAHRFKDLGGENFKALVLIQVLAQVLOQCPFEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAAKAACLLP 180
Db 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCCAAKAACLLP 204
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPKFAEFAVSKLVDLT 240
Db 205 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPKFAEFAVSKLVDLT 264
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDISSKLKECCCKPILLEKSHCHIAEVNDEMPA 300
Db 265 VHTCCCHGDLLECCADDRADLAKYICENQDISSKLKECCCKPILLEKSHCHIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYVARRHPDYSVVLLRLAKTVEITLLEK 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYVARRHPDYSVVLLRLAKTVEITLLEK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGSEYKFNALLVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGSEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPKAPKPCADYLSVLNQLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRLGKVGSKCKHPKAPKPCADYLSVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 564
QY 541 KEOLKAVMDDDFAAFVEKCKKADDETCFAEEGKKLVAAQAALGLMSPRLL----- 590
Db 565 KEOLKAVMDDDFAAFVEKCKKADDETCFAEEGKKLVAAQAALGLAPPRLICDSRVLYRY 624
QY 591 -----EVPSCSHA-----LPQGLS 603
Db 625 LLEAKEAEQITTCGAHCSSLNEQITVPDTKVNFWYKMEVEGQQAQVEMVQGLLSEAVL 684
QY 604 PGQVIVIRGLVLOEP-----KHTVSLRDOQAHAHVTLRASPADRTLQWISRW 651
Db 685 RGQALLVQSSQPWEPLQLQHVHDKAVGLRSLTLLR-----ALRA----- 723
QY 652 GQKKLISAP-----FLFYPRFFEVLLLFQEGGKL 682
```

724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYGNFLRGKLUK 764

RESULT 15

ADFL14959  
ID ADF14959 standard; protein; 774 AA.

AC ADF14959;

DT 12-FEB-2004 (first entry)

DE Human albumin therapeutic fusion protein SeqID255.

XX albumin fusion protein; albumin activity; human serum albumin;  
XX serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.

OS Chimeric.

OS Homo sapiens.

XX W02003060071-A2.

XX 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US040891.

XX 21-DEC-2001; 2001US-0341811P.

XX 28-JAN-2002; 2002US-0350358P.

XX 26-FEB-2002; 2002US-0359370P.

XX 27-MAR-2002; 2002US-0367500P.

XX 08-APR-2002; 2002US-0370227P.

XX 10-MAY-2002; 2002US-0378950P.

XX 24-MAY-2002; 2002US-0382617P.

XX 05-JUN-2002; 2002US-0385708P.

XX 24-JUL-2002; 2002US-0398008P.

XX 09-AUG-2002; 2002US-0402131P.

XX 13-AUG-2002; 2002US-0402708P.

XX 18-SEP-2002; 2002US-0411355P.

XX 02-OCT-2002; 2002US-0414984P.

XX 11-OCT-2002; 2002US-0417611P.

XX 23-OCT-2002; 2002US-0420246P.

XX 05-NOV-2002; 2002US-0423623P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

XX (PRIN-) PRINCIPIA PHARM CORP.

XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;

XX WPI; 2003-598517/56.

XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.

XX Example 4; SEQ ID NO 255; 24pp; English.

XX This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
XX significant proportion of the osmotic pressure of serum and also  
XX functions as a carrier of endogenous and exogenous ligands. The fusion of  
XX albumin to a therapeutic protein may increase shelf-life and stability of  
XX the therapeutic protein. The albumin fusion protein of the invention may  
XX allow production of compositions with antidiabetic activity whilst the  
XX nucleotide sequence which encodes it may be useful for gene therapy. The  
XX albumin fusion protein is useful for preparing a composition for treating  
XX diabetes mellitus. The present sequence is the amino acid sequence of a  
XX novel full-length human albumin therapeutic fusion protein of the

CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

XX Sequence 774 AA;

Query Match 83.2%; Score 3151; DB 7; Length 774;  
Best local Similarity 80.5%; Pred. No. 7.1e-73;  
Matches 617; Conservative 9; Mismatches 30; Indels 110; Gaps 8;

Qy 1 DAHKSEVAHFRKDLGRENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADSEAE 60

Db 25 DAHKSEVAHFRKDLGRENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADSEAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEV 120

Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEV 144

Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180

Db 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204

Qy 181 KLDELDRDEGRKASSAKQRLKCSLQKGERAFKAWAVARLSORFPKABFAEVSKLVTDLTK 240

Db 205 KLDELDRDEGRKASSAKQRLKCSLQKGERAFKAWAVARLSORFPKABFAEVSKLVTDLTK 264

Qy 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300

Db 265 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 360

Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLK 384

Qy 361 CAAADPHECYAKVDFEFPKLVVEEPQNLIKONCELFEOQGEYKFNALLVRYTKKVPQVST 420

Db 385 CAAADPHECYAKVDFEFPKLVVEEPQNLIKONCELFEOQGEYKFNALLVRYTKKVPQVST 444

Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKRMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480

Db 445 PTLVEVSRNLGKVGSKCKHPKAEKRMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504

Qy 481 LVNRRPCFSALEVDVETVVPKFNATETTFHADICTLSEKEROIKKQATLVELVHKHPKAT 540

Db 505 LVNRRPCFSALEVDVETVVPKFNATETTFHADICTLSEKEROIKKQATLVELVHKHPKAT 564

Qy 541 KEQLKAVMDDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPR----- 589

Db 565 KEQLKAVMDDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPR----- 624

Qy 590 -LEVPCSHALP-----OGLS----- 603

Db 625 LLEAKEAEATITGCAEHCSLNEALITVPTDKYFNFKMEVGGQAVEVWQGLALLSEAVL 684

Qy 604 PGQVIVRGLVLQEP-----KHFTVSRDQAAHAPVTLIRASFADRTLOWISRW 651

Db 685 RGQALLVASSQPWEPLQLHVDKAVSGLRSLTLRL-----ALRA----- 723

Qy 652 GQKLLISAP-----FLFVPORFREVLILLFQSGGLKL 682

Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYGNFLRGKLUK 764

RESULT 16

ADFL15004  
ID ADF15004 standard; protein; 774 AA.

XX ADF15004;

XX DT 12-FEB-2004 (first entry)

XX Human albumin therapeutic fusion protein SeqID300.

XX

KW albumin fusion protein; albumin activity; human serum albumin;  
 KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
 KW gene therapy; diabetes mellitus; human.

XX Chimeric.

OS Homo sapiens.

XX WO2003060071-A2.

XX 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US040891.

XX 21-DEC-2001; 2001US-0341811P.

XX 24-JAN-2002; 2002US-0350358P.

XX 28-JAN-2002; 2002US-0351360P.

XX 26-FEB-2002; 2002US-0359370P.

XX 28-FEB-2002; 2002US-0360000P.

XX 27-MAR-2002; 2002US-0367500P.

XX 08-APR-2002; 2002US-0370227P.

XX 10-MAY-2002; 2002US-0378950P.

XX 24-MAY-2002; 2002US-0382617P.

XX 28-MAY-2002; 2002US-0383123P.

XX 05-JUN-2002; 2002US-0385708P.

XX 10-JUL-2002; 2002US-0394625P.

XX 24-JUL-2002; 2002US-0398008P.

XX 09-AUG-2002; 2002US-0402131P.

XX 13-AUG-2002; 2002US-0402708P.

XX 18-SEP-2002; 2002US-0411355P.

XX 18-SEP-2002; 2002US-0411426P.

XX 02-OCT-2002; 2002US-0414984P.

XX 11-OCT-2002; 2002US-0417611P.

XX 23-OCT-2002; 2002US-0420246P.

XX 05-NOV-2002; 2002US-0423623P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (DELZ) DELTA BIOTECHNOLOGY LTD.

PA (PRIN-) PRINCIPIA PHARM CORP.

XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;

XX WPI; 2003-598517/56.

XX New albumin fusion protein, useful for preparing a composition for

XX treating diabetes mellitus.

XX Example 4; SEQ ID NO 300; 24pp; English.

XX This invention relates to a novel albumin fusion protein having albumin

XX or biological activity. Human serum albumin is responsible for a

XX significant proportion of the osmotic pressure of serum and also

XX functions as a carrier of endogenous and exogenous ligands. The fusion of

XX albumin to a therapeutic protein may increase shelf-life and stability of

XX the therapeutic protein. The albumin fusion protein of the invention may

XX allow production of compositions with antidiabetic activity whilst the

XX nucleotide sequence which encodes it may be useful for gene therapy. The

XX albumin fusion protein is useful for preparing a composition for treating

XX diabetes mellitus. The present sequence is the amino acid sequence of a

XX novel full-length human albumin therapeutic fusion protein of the

XX invention. Note: The sequence data for this patent did not form part of

XX the printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

XX

XX Query Match 83.2%; Score 3151; DB 7; Length 774;

XX Best Local Similarity 80.3%; Pred. No. 7.1e-73;

XX Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;

XX

XX 1 DAHKSEVAHRFKDILGEENFKALVLIJAFQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 60

XX 25 DAHKSEVAHRFKDILGEENFKALVLIJAFQYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 180

DB 145 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 204

QY 181 KLDELDEGKASSAKORLAKCASLOKGERAFKAWAVARLSORPFAEFKAEVSKLVTDLT 240

DB 205 KLDELDEGKASSAKORLAKCASLOKGERAFKAWAVARLSORPFAEFKAEVSKLVTDLT 264

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDISISSKKECCCKPILLEKSHCIAEVENDEMPA 300

DB 265 VHTCCCHGDLLECCADDRADLAKYICENQDISISSKKECCCKPILLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLLRLLAKTYETTLEKC 360

DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVLLLRLLAKTYETTLEKC 384

QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420

DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVNLQLCVLHEKTPVSDRVTCKCTES 480

DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVNLQLCVLHEKTPVSDRVTCKCTES 504

QY 481 LVNRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540

DB 505 LVNRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKKLVAAASQAALGLMSPLI----- 590

DB 565 KEQLKAVMDDDFAAFVEKCKCKADDKETCFAEEGKKLVAAASQAALGLMSPLI----- 624

QY 591 -----EVPCSHA-----LPOGLS 603

DB 625 LLEAKEAENITTCGAHCSLNENITVDTKVNFYANKMEVGGQAVEVWQGLALLSEAVL 684

QY 604 PGQVIIVRGLVLOEP-----KHFTVSLRDOAAHAPVTLRASPADRTLQWISRW 651

DB 685 RGQALLVNSSQPWEPQLQHVDAVSGLSLTLTLR-----ALRA----- 723

QY 652 GQKLLSAP-----FLFYQRFVEVLLLFQEGGK 682

DB 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFRUGK 764

# RESULT 17

ADFI4949  
 ID ADFI4949 standard; protein; 774 AA.

XX ADFI4949;

XX 12-FEB-2004 (first entry)

XX Human albumin therapeutic fusion protein SeqID245.

XX albumin fusion protein; albumin activity; human serum albumin;  
 KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
 KW gene therapy; diabetes mellitus; human.

XX Chimeric.

OS Homo sapiens.

XX WO2003060071-A2.

XX 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US040891.

XX 21-DEC-2001; 2001US-0341811P.

24-JAN-2002; 2002US-0350358P.  
 28-JAN-2002; 2002US-0351360P.  
 26-FEB-2002; 2002US-0359370P.  
 28-FEB-2002; 2002US-0360000P.  
 27-MAR-2002; 2002US-0367500P.  
 08-APR-2002; 2002US-0370227P.  
 10-MAY-2002; 2002US-0378950P.  
 24-MAY-2002; 2002US-0382617P.  
 05-JUN-2002; 2002US-0385708P.  
 10-JUL-2002; 2002US-0394625P.  
 24-JUL-2002; 2002US-0398088P.  
 09-AUG-2002; 2002US-0402131P.  
 13-AUG-2002; 2002US-0402708P.  
 18-SEP-2002; 2002US-0411355P.  
 18-SEP-2002; 2002US-0411426P.  
 02-OCT-2002; 2002US-0414984P.  
 11-OCT-2002; 2002US-0417611P.  
 23-OCT-2002; 2002US-0420246P.  
 05-NOV-2002; 2002US-0423623P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 (DELZ) DELTA BIOTECHNOLOGY LTD.  
 (PRIN-) PRINCIPIA PHARM CORP.  
 Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
 WPI; 2003-598517/56.  
 New albumin fusion protein, useful for preparing a composition for  
 treating diabetes mellitus.  
 Example 4; SEQ ID NO 245; 24pp; English.  
 This invention relates to a novel albumin fusion protein having albumin  
 or biological activity. Human serum albumin is responsible for a  
 significant proportion of the osmotic pressure of serum and also  
 functions as a carrier of endogenous and exogenous ligands. The fusion of  
 albumin to a therapeutic protein may increase shelf-life and stability of  
 the therapeutic protein. The albumin fusion protein of the invention may  
 allow production of compositions with antidiabetic activity whilst the  
 nucleotide sequence which encodes it may be useful for gene therapy. The  
 albumin fusion protein is useful for preparing a composition for treating  
 diabetes mellitus. The present sequence is the amino acid sequence of a  
 novel full-length human albumin therapeutic fusion protein of the  
 invention. Note: The sequence data for this patent did not form part of  
 the printed specification, but was obtained in electronic format directly  
 from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
 Sequence 774 AA;  
 Query Match 83.2%; Score 3151; DB 7; Length 774;  
 Best Local Similarity 80.3%; Pred. No. 7.1e-73;  
 Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;  
 1 DAHSEVAHRFKDLGEENFKALVLIAPYQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60  
 25 DAHSEVAHRFKDLGEENFKALVLIAPYQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 84  
 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 120  
 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVRPEV 144  
 121 DVMCTAFHNEFTFLKKLYLETARHPHYAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180  
 145 DVMCTAFHNEFTFLKKLYLETARHPHYAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 204  
 181 KLDELURDEKASSAKQRLKASLQKFGERAFKAWAVARLSQRFKAEFAEYVKLVTDLT 240  
 205 KLDELURDEKASSAKQRLKASLQKFGERAFKAWAVARLSQRFKAEFAEYVKLVTDLT 264  
 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
 Qy 301 DLPFLAADFVESKDVCKNRYAEAKDVFGLMFLYFYARRHPDYSVVLLRLAKTYTTLEK 360  
 Db 325 DLPFLAADFVESKDVCKNRYAEAKDVFGLMFLYFYARRHPDYSVVLLRLAKTYTTLEK 384  
 Qy 361 CAAADPHECYAKVDFEPKPLVESPNLIKONCELFEOLGEYKFNALLVRYTKVPOVST 420  
 Db 385 CAAADPHECYAKVDFEPKPLVESPNLIKONCELFEOLGEYKFNALLVRYTKVPOVST 444  
 Qy 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 Db 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504  
 Qy 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
 Db 505 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKEROIKKQATLVELVGHKPKAT 564  
 Qy 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFABEGKLVAAASQAALGMSPRLL 590  
 Db 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFABEGKLVAAASQAALGMSPRLL 624  
 Qy 591 -----EVPCSHA-----LPOGLS 603  
 Db 625 LLEAKEAENITTCGAHCSLNENITVDTKVNFYAKRMEVGOQAVEVWQGLALLSEAVL 684  
 Qy 604 PGQVLIIVRLVQLQEP-----KHFTVSLRDOAAHAPVTLTRASFADRTLQWISRW 651  
 Db 685 RGQALLVNSQWPEFLQHLVHDKAVSGLSLTLLR-----ALRA----- 723  
 Qy 652 GQKLLISAP-----FLFYQPQFFEVLLLFQEGGLKL 682  
 Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYGNFLRGKLU 764  
 RESULT 18  
 ADF14960  
 ID ADF14960 standard; protein; 774 AA.  
 XX  
 AC ADF14960;  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Human albumin therapeutic fusion protein SeqID256.  
 KW albumin fusion protein; albumin activity; human serum albumin;  
 KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
 KW gene therapy; diabetes mellitus; human.  
 XX  
 OS Chimeric.  
 OS Homo sapiens.  
 XX  
 PN WO2003060071-A2.  
 XX  
 PD 24-JUL-2003.  
 XX  
 XX 23-DEC-2002; 2002WO-US040891.  
 XX  
 XX 21-DEC-2001; 2001US-0341811P.  
 XX 24-JAN-2002; 2002US-0350358P.  
 XX 28-JAN-2002; 2002US-0351360P.  
 XX 26-FEB-2002; 2002US-0359370P.  
 XX 28-FEB-2002; 2002US-0360000P.  
 XX 27-MAR-2002; 2002US-0367500P.  
 XX 08-APR-2002; 2002US-0370227P.  
 XX 10-MAY-2002; 2002US-0378950P.  
 XX 24-MAY-2002; 2002US-0382617P.  
 XX 28-MAY-2002; 2002US-0385708P.  
 XX 05-JUN-2002; 2002US-0394625P.  
 XX 24-JUL-2002; 2002US-0398088P.  
 XX 09-AUG-2002; 2002US-0402131P.  
 XX 13-AUG-2002; 2002US-0402708P.

PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (DEL2 ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX  
XX WPI; 2003-598517/56.  
XX  
PT New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX  
XX Example 4; SEQ ID NO 256; 24pp; English.  
XX  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
SQ Sequence 774 AA;

Query Match 83.28; Score 3151; DB 7; Length 774;  
Best Local Similarity 80.38; Pred. No. 7.1e-73;  
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;  
  
QY 1 DAHKSEVAHFKD LGNEFKALVLIAPAOYLQCPPEFHVKNVEVTEFAKTCVADESAE 60  
DB 25 DAHKSEVAHFKD LGNEFKALVLIAPAOYLQCPPEFHVKNVEVTEFAKTCVADESAE 84  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
  
QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFEAKRYKAFTCCOAAADKAAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFEAKRYKAFTCCOAAADKAAACLLP 204  
  
QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRFPKAEFAEVSKLVDLTJK 240  
DB 205 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRFPKAEFAEVSKLVDLTJK 264  
  
QY 241 VHTCCCHGDLLECAADRADLAKYICENQDSISSKLKECCCKPLLEKSHCHIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECAADRADLAKYICENQDSISSKLKECCCKPLLEKSHCHIAEVENDEMPA 324  
  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTVETTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVVLLRLAKTVETTLEKC 384  
  
QY 361 CAAADPHCEYAKVFDEPKPLVEEPQNLIKNCLEFEQKGEYKFKONALLVRYTKKVPQVST 420  
DB 385 CAAADPHCEYAKVFDEPKPLVEEPQNLIKNCLEFEQKGEYKFKONALLVRYTKKVPQVST 444  
  
QY 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPKAEADYLSVLNQLCVLHKEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKKHPKAEKMPKAEADYLSVLNQLCVLHKEKTPVSDRVTKCCTES 504

QY 481 LVNRRPFCFSALEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540  
DB 505 LVNRRPFCFSALEVDYTVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 564  
  
QY 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGLMSPRI----- 590  
DB 565 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGLMSPRI----- 624  
  
QY 591 -----EYPCSHA-----LPQGLS 603  
DB 625 LLEAKEAENITTCABHCSLNENITVPTKVNFYAKRMEVGOQAAVEVWQGLALLSEAVL 684  
  
QY 604 PGQVIIVRGVLQEP-----KHFTVSLRDQAAHAPVTLRASFADRTLQWISRW 651  
DB 685 RGQALLVNSQPEWPLQHLVDKAVSGLSITLLR-----ALRA----- 723  
  
QY 652 GQKKLISAP-----FLFYQRPFEVLLLFQEGGLKL 682  
DB 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFLRGKGLKL 764  
  
RESULT 19  
ADFI4923  
ID ADFI4923 standard; protein; 868 AA.  
XX  
AC ADFI4923;  
XX  
DT 12-FEB-2004 (first entry)  
DE  
DE Human albumin therapeutic fusion protein SeqID219.  
XX  
XX albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX  
OS Chimeric.  
OS Homo sapiens.  
XX  
XX WO2003060071-A2.  
XX  
PD 24-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US040891.  
XX  
XX 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350558P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DEL2 ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX  
XX WPI; 2003-598517/56.  
XX

XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PS Example 4; SEQ ID NO 219; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/publishedpt\_sequences  
XX Sequence 868 AA;  
SQ

Query Match 83.2%; Score 3151; DB 7; Length 868;  
Best Local Similarity 70.5%; Pred. No. 8.6e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
QY 1 DAHSEVAHPRKDI GEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
DB 25 DAHSEVAHPRKDI GEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 144  
QY 121 DVMTAFHNDNETFLKYLBIARRHPYFPAPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 145 DVMTAFHNDNETFLKYLBIARRHPYFPAPELLFFAKRYKAAFTCCQAAADKAACLLP 204  
QY 181 KLDELREGKASSAKORLKCSLQKFGERAFAKAVARLSORFPAEVS KLVTDLTK 240  
DB 205 KLDELREGKASSAKORLKCSLQKFGERAFAKAVARLSORFPAEVS KLVTDLTK 264  
QY 241 VHTCCGDLLECAADDRADLAKYCENQDSISSKLKCECKPFLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCGDLLECAADDRADLAKYCENQDSISSKLKCECKPFLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKQVCKNYABAKOVFLGMFLYEVARRHPDYSVLLLR LAKTYETTLK 360  
DB 325 DLPSLAADFVSKQVCKNYABAKOVFLGMFLYEVARRHPDYSVLLLR LAKTYETTLK 384  
QY 361 CAADPHCEYAKVDFEPFLVEEPQNLIKONCELPQOLGEYKFNALLAVRYTKKVPQVST 420  
DB 385 CAADPHCEYAKVDFEPFLVEEPQNLIKONCELPQOLGEYKFNALLAVRYTKKVPQVST 444  
QY 421 PTLVEVSNLGVSKCKCHPEAKMPCAEDYLSVNLQLCVLEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSNLGVSKCKCHPEAKMPCAEDYLSVNLQLCVLEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYTPVPEFNAETFTTFHADICTLSEKROIKKQATLVELVHKHKPAT 540  
DB 505 LVNRRPCFSALEVDYTPVPEFNAETFTTFHADICTLSEKROIKKQATLVELVHKHKPAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGMSRLVLP----- 593  
DB 565 KEQLKAVMDDDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGVA---ETPTYPWRDA 621  
QY 594 -----CSHALPQGLSPQVIVRGLVLOE-----PKHFT----- 622  
DB 622 ETGERLVCAQCP-----GTFVQRCRRDSTTCPCPFRHYTQFVNYLCR 669  
QY 623 -----VSLRDOAAHA----- 632

Db 670 YCNVLGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGTSPQNTQ 729  
QY 633 -----PVTLRAS----- 639  
Db 730 CQPCPCTFSASSSSSESQCPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRVPGA 789  
QY 640 -----PADRTLOWISR-----WG-----OKKLISAPFLFYPPQRF 668  
Db 790 ECEBRAVIDFVAFQDISIKRLQRLLOALEAPEGPGTTPRAGRAALQLKL-----RRR 840  
QY 669 FEVLLLFQEGGLKALNGQGLGATSMNQQALEQLRELRIISG 709  
Db 841 LTELHGQDQALLVRL-----LQALRVARMPG 867  
RESULT 20  
ADH21275  
ID ADH21275 standard; protein; 868 AA.  
XX AC ADH21275;  
XX DT 11-MAR-2004 (first entry)  
XX Human albumin/TF6 fragment fusion protein, SEQ ID NO:72.  
XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
KW shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;  
KW anorectic; ophthalmological; gene therapy.  
XX Chimeric.  
OS Homo sapiens.  
XX WO2003059934-A2.  
XX 24-JUL-2003.  
XX 23-DEC-2002; 2002WO-US040892.  
XX 21-DEC-2001; 2001US-0341811P.  
XX 24-JAN-2002; 2002US-0350358P.  
XX 26-FEB-2002; 2002US-0359370P.  
XX 28-FEB-2002; 2002US-0360000P.  
XX 27-MAR-2002; 2002US-0367500P.  
XX 08-APR-2002; 2002US-0370227P.  
XX 10-MAY-2002; 2002US-0378950P.  
XX 24-JUL-2002; 2002US-0398008P.  
XX 09-AUG-2002; 2002US-0402131P.  
XX 13-AUG-2002; 2002US-0402708P.  
XX 18-SEP-2002; 2002US-0411355P.  
XX 02-OCT-2002; 2002US-0414984P.  
XX 11-OCT-2002; 2002US-0417611P.  
XX 23-OCT-2002; 2002US-0420246P.  
XX 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Haseltine WA;  
XX WPI; 2003-598501/56.  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PS Disclosure; SEQ ID NO 72; 1086pp; English.  
XX The invention relates to fusion proteins comprising human serum albumin  
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments. The therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini of

CC albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein, the method of extending the shelf-life of a  
CC therapeutic protein by fusion with albumin, and the treatment of disease  
CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hypoglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity. The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.  
XX  
SQ Sequence 868 AA;

Query Match 83.2%; Score 3151; DB 7; Length 868;  
Best Local Similarity 70.5%; Pred. No. 8.6e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHKSEVAHFCDLGFENFKALVLIAPAOYLQCPFEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHKSEVAHFCDLGFENFKALVLIAPAOYLQCPFEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAAACLLP 204  
QY 181 KLDELDEGKASAKQRLKASQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTQ 240  
DB 205 KLDELDEGKASAKQRLKASQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTQ 264  
QY 241 VHTECCHGDLLECCADRADLAKYICENQDISSKLKECCCKPLEKSHCIAEVENDEMPA 300  
DB 265 VHTECCHGDLLECCADRADLAKYICENQDISSKLKECCCKPLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYERARHPDYSVVLRLAKTYETTTLEK 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYERARHPDYSVVLRLAKTYETTTLEK 384  
QY 361 CAADPHECYAKVDFEKPFLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKVPQVST 420  
DB 385 CAADPHECYAKVDFEKPFLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCABDYLSSVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 445 PTLVEVSRLNGKVGSKCKHPEAKRMPCABDYLSSVLNQLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDVETVPKFEFAETFTFHADICTLSEKERQIKQATALVELVKHPKAT 540  
DB 505 LVNRRPCFSALEVDVETVPKFEFAETFTFHADICTLSEKERQIKQATALVELVKHPKAT 564  
QY 541 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMSPRLEVP----- 593  
DB 565 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLVA---ETPTYPWRDA 621  
QY 594 -----CShALPQGLSPGVIIIVRLVLOE-----PKHFT----- 622  
DB 622 ETGERLVCAQCCP-----GTFVQPCRRDSTPTTCGPPPHRYTFQWNYLERCR 669  
QY 623 -----VSLRQOAHVA----- 632  
DB 670 YCNVLCGERBEEARACHATHNRACRGTGFFAHAGFCLCHASCPGAGVIAPGTPSQNTQ 729  
QY 633 -----PVTLRAS----- 639

DB 730 CQPCPGTFSASSSSSEQCPHRCNTALGLALNVPGSSSHDTLTCTGFFLSTRVPGA 789  
QY 640 -----FADRTLQWISR-----WG-----QKLLISAPFLFVPORF 668  
DB 790 ECERAVIDFVAFQDISIKRLQRLQLQALEAPEGCGPTFRAGRALQLKL-----RRR 840  
QY 669 FEVLILFOGGGKLALNGOGLGATSNMQALEQLRELISG 709  
DB 841 LTELIGAQDALLVRL-----LQALRVARMPG 867

RESULT 21  
ADF14922  
ID ADF14922 standard; protein; 874 AA.  
XX  
AC ADF14922;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Human albumin therapeutic fusion protein SeqID218.  
XX  
KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
XX  
OS Chimeric.  
OS Homo sapiens.  
XX  
PN WO2003060071-A2.  
XX  
PD 24-JUL-2003.  
XX  
PF 23-DEC-2002; 2002WO-US040891.  
XX  
PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 28-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX  
XX WPI; 2003-598517/56.  
DR  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX  
XX Example 4; SEQ ID NO 218; 24pp; English.  
PS  
XX  
CC This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also



functions as a carrier of endogenous and exogenous ligands. The fusion of albumin to a therapeutic protein may increase shelf-life and stability of the therapeutic protein. The albumin fusion protein of the invention may allow production of compositions with antidiabetic activity whilst the nucleotide sequence which encodes it may be useful for gene therapy. The albumin fusion protein is useful for preparing a composition for treating diabetes mellitus. The present sequence is the amino acid sequence of a novel full-length human albumin therapeutic fusion protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

Query Match 83.2%; Score 3151; DB 7; Length 874;  
Best Local Similarity 70.5%; Pred. No. 8.7e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAKHSEVAHRFKDILGENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 25 DAKHSEVAHRFKDILGENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLETETGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLETETGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 144  
Qy 121 DVMCTAFHDNEETFLKYLVIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKYLVIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPAKFAEVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFPAKFAEVSKLVTDLTK 264  
Qy 241 VHTCCGGDLLECCADRDADLAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGGDLLECCADRDADLAKYICENQDSISSKLKCECKPLEKSHCIAEVENDEMPA 324  
Qy 301 DLPLSADPFVSKDVKRYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360  
Db 325 DLPLSADPFVSKDVKRYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384  
Qy 361 CAAADPHECYAKVDFEKPVEEQLNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFEKPVEEQLNLIKONCELFQELGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRLGKSGKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKSGKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 564  
Qy 541 KEQLKAVNMDPFAAFVEKCKKADDETCFAEKGKLVAAASQAALGMLSPRLVPP 593  
Db 565 KEQLKAVNMDPFAAFVEKCKKADDETCFAEKGKLVAAASQAALGLVA---ETPTVPMRDA 621  
Qy 594 -----CSHALPQGLSPGVIIIVGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCAQCP-----GTFVQPCRRDSTTCGCPPRHYTQFWNYLRCR 669  
Qy 623 -----VSLRDQAAHA----- 632  
Db 670 YCNVLGREGREEARACHATHNACRGTGFFAHAGFLEHASCPCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 QCPCPPTGFSASSSSSEQCPHRCNCTALGLALNVPSSSHDITLCTSGTGFPLSTRVPGA 789  
Qy 640 -----PADTLQWISR-----WG-----QKLLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFODISIKRLQRLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840

Qy 669 FEVLLIFQEGGLKALNQGLGATSMNQQAQLRELRLISG 709  
Db 841 LTELGAQDQALLVRL-----LQALRVARMFG 867

RESULT 22

ADH21274  
ID ADH21274 standard; protein; 874 AA.  
AC ADH21274;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Human albumin/TR6 fragment fusion protein, SEQ ID NO:71.

Fusion protein; human serum albumin; HSA; therapeutic protein; shelf-life; in vitro biological activity; in vivo biological activity; metabolic disorder; endocrine disorder; diabetes; type 1; type 2; diabetes-related condition; hyperglycaemia; neural disorder; neuropathy; retinopathy; cardiovascular disorder; heart disease; renal disorder; obesity; glucose level maintenance; weight loss; antidiabetic; cardiatic; anorectic; ophthalmological; gene therapy.

Chimeric.  
OS Homo sapiens.

XX WO2003059934-A2.

PN 24-JUL-2003.

PD 23-DEC-2002; 2002WO-US040892.

XX 21-DEC-2001; 2001US-0341811P.

PR 24-JAN-2002; 2002US-0350358P.

PR 26-FEB-2002; 2002US-0359370P.

PR 28-FEB-2002; 2002US-0360000P.

PR 27-MAR-2002; 2002US-0367500P.

PR 08-APR-2002; 2002US-0370227P.

PR 10-MAY-2002; 2002US-0378950P.

PR 24-JUL-2002; 2002US-0402131P.

PR 09-AUG-2002; 2002US-0402708P.

PR 13-AUG-2002; 2002US-0411355P.

PR 18-SEP-2002; 2002US-0414984P.

PR 02-OCT-2002; 2002US-0417611P.

PR 11-OCT-2002; 2002US-0420246P.

PR 23-OCT-2002; 2002US-0423623P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Haseltine WA;

XX WPI; 2003-598501/56.

DR New albumin fusion protein, useful for preparing a composition for treating diabetes mellitus.

XX Disclosure; SEQ ID NO 71; 1086pp; English.

The invention relates to fusion proteins comprising human serum albumin (ADH21530) and a therapeutic polypeptide such as a therapeutic protein, antibody or peptide or their variants or fragments. The therapeutic protein may be fused to the N-terminus, the C-terminus or both termini of albumin via a linker. The albumin component of the fusion proteins prolongs the shelf-life and the in vitro and vivo biological activity of the proteins compared with those of the corresponding therapeutic proteins on their own. The invention also relates to nucleic acids encoding albumin fusion proteins, vectors and host cells comprising an albumin fusion protein nucleic acid, compositions and kits comprising an albumin fusion protein, the method of extending the shelf-life of a therapeutic protein by fusion with albumin, and the treatment of disease using an albumin fusion protein. The albumin fusion proteins may be used

CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease, renal disorders and obesity. The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.  
XX  
SQ Sequence 874 AA;  
Query Match 83.2%; Score 3151; DB 7; Length 874;  
Best Local Similarity 70.5%; Pred. No. 8.7e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
QY 1 DAHSEVAHRFDLGEENFKALVLIAPQAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAB 60  
DB 25 DAHSEVAHRFDLGEENFKALVLIAPQAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAB 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144  
QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
DB 145 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLOKGFERAFKAWAVARLSQRPKFAEFAVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLOKGFERAFKAWAVARLSQRPKFAEFAVSKLVTDLTK 264  
QY 241 VHTTECHGDLLECCADRADLAKYICENQDISSKLKECEKPLELXSHCIAEVENDDEMA 300  
DB 265 VHTTECHGDLLECCADRADLAKYICENQDISSKLKECEKPLELXSHCIAEVENDDEMA 324  
QY 301 DLPSLAADFVESKVCNVAEAKDFLGMFLYIYARRHPDYSVVLLRLRAKTYETTLK 360  
DB 325 DLPSLAADFVESKVCNVAEAKDFLGMFLYIYARRHPDYSVVLLRLRAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420  
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPYSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPYSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKHKPKAT 540  
DB 505 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKHKPKAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKLVAAQALGLMSPLREVP----- 593  
DB 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKLVAAQALGLVA---ETPTPWRDA 621  
QY 594 -----CSHALPQGLSPGQVIVRGLVQS-----PKHFT----- 622  
DB 622 ETGERLVCAQCP-----GTFVQRPCRDSPTTCGCPPRHYTFQWNYLCR 669  
QY 623 -----VSLRDOAAHA----- 632  
DB 670 YCNVLGEREEERACHATHNACRGTGFFAHAGFCLEHASCPPGAGVIAPGTPSQNTQ 729  
QY 633 -----PVTLRAS----- 639  
DB 730 CQPCPPTFSASSSSSEQCPQRNCTALGLNALVPGSSSHDTLCTSGTGFPLSTRVPGAE 789  
QY 640 -----FADRLQWISR-----WG-----QKLLISAPFLFPQRF 668  
DB 790 ECERAVIDVFVAFQDISIKRLQRLLOALEAPGEGGTPFRAGRAALQKL-----RRR 840  
QY 669 FEVLLLFQEGGLKALNGQGLGATSMNQQAOLEQRLRLISG 709

DB 841 LTELGAQDQGALLVRL-----LQALRVARMPG 867  
RESULT 23  
ADF14925  
ID ADF14925 standard; protein; 880 AA.  
XX  
AC ADF14925;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Human albumin therapeutic fusion protein SeqID221.  
XX  
KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
KW gene therapy; diabetes mellitus; human.  
OS Chimeric.  
OS Homo sapiens.  
XX  
PN WO2003060071-A2.  
XX  
PD 24-JUL-2003.  
XX  
PF 23-DEC-2002; 2002WO-US040891.  
XX  
PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX  
PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
DR  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
PT  
XX  
PS Example 4; SEQ ID NO 221; 24pp; English.  
XX  
CC This invention relates to a novel albumin fusion protein having albumin  
CC or biological activity. Human serum albumin is responsible for a  
CC significant proportion of the osmotic pressure of serum and also  
CC functions as a carrier of endogenous and exogenous ligands. The fusion of  
CC albumin to a therapeutic protein may increase shelf-life and stability of  
CC the therapeutic protein. The albumin fusion protein of the invention may  
CC allow production of compositions with antidiabetic activity whilst the  
CC nucleotide sequence which encodes it may be useful for gene therapy. The  
CC albumin fusion protein is useful for preparing a composition for treating  
CC diabetes mellitus. The present sequence is the amino acid sequence of a  
CC novel full-length human albumin therapeutic fusion protein of the  
CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
SQ Sequence 880 AA;

Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHSEVAHRFKDIBENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60  
Db 25 DAHSEVAHRFKDIBENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVMTAFHDNEBETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMTAFHDNEBETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204  
Qy 181 KLDELREGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAFVSKLVTLDTLK 240  
Db 205 KLDELREGKASSAKORLKCASLOKFGERAFAKAVARLSORFPKAFVSKLVTLDTLK 264  
Qy 241 VHTCCCHGDLLECADDRLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECADDRLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVBSKOVCKNYAEAKDVLGMFLIYVARRHPDYSVVLLRLAKTYETTLEKC 360  
Db 325 DLPSLAADFVBSKOVCKNYAEAKDVLGMFLIYVARRHPDYSVVLLRLAKTYETTLEKC 384  
Qy 361 CAAADPHCYAKVDFDKPLVEEONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHCYAKVDFDKPLVEEONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRLGKVGSKCCKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCCKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYKPEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 564  
Qy 541 KEQLKAVNMDFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPLREVP----- 593  
Db 565 KEQLKAVNMDFAAFVEKCKADDKETCFABEGKKLVAASQAALGLVA---ETPTYPWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIIRGLVLOE-----PKHFT----- 622  
Db 622 ETGERLVCACQPP-----GTFVQRPCCRDSPTTCGCPPRHYTQFWNYLERCR 669  
Qy 623 -----VSLRDQAAHA----- 632  
Db 670 YCNVLGCEBEARACHATHNRCRGTGFFAHAGFCLEHASCPCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 QCPCPPTGTFSSSSSSSEQQPHRNCTALGLALNVPSSSHDTLCTSGTGFPLSTRVPGA 789  
Qy 640 -----FADRTLQWISR-----WG-----QKLLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQLLQALRAPEGWGTPPAGRALQLKL-----RRR 840  
Qy 669 FVLLLFQEGGKLKALNGOGLGATSNQOALFQELRLRISG 709  
Db 841 LTELIGAQQDGLLVRL-----LQALVARMPG 867

RESULT 24  
ADF14983  
ID ADF14983 standard; protein; 880 AA.

XX ADF14983;  
XX DT 12-FEB-2004 (first entry)  
XX DE Human albumin therapeutic fusion protein SeqID279.  
XX KW albumin fusion protein; albumin activity; human serum albumin;  
KW serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.  
XX OS Chimeric.  
XX OS Homo sapiens.  
XX WO2003060071-A2.  
XX PD 24-JUL-2003.  
XX PF 23-DEC-2002; 2002WO-US040891.  
XX PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 28-JAN-2002; 2002US-0351360P.  
PR 28-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
PA (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX Example 4; SEQ ID NO 279; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
XX significant proportion of the osmotic pressure of serum and also  
XX functions as a carrier of endogenous and exogenous ligands. The fusion of  
XX albumin to a therapeutic protein may increase shelf-life and stability of  
XX the therapeutic protein. The albumin fusion protein of the invention may  
XX allow production of compositions with antidiabetic activity whilst the  
XX nucleotide sequence which encodes it may be useful for gene therapy. The  
XX albumin fusion protein is useful for preparing a composition for treating  
XX diabetes mellitus. The present sequence is the amino acid sequence of a  
XX novel full-length human albumin therapeutic fusion protein of the  
XX invention. Note: The sequence data for this patent did not form part of  
XX the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX SQ Sequence 880 AA;

Query Match 83.2%; Score 3151; DB 7; Length 880;

Best Local Similarity 70.5%; Pred. No. 8.8e-73;		Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;	
Qy	1	DAHKSEVAHRFKDGLGENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADES	60
Db	25	DAHKSEVAHRFKDGLGENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADES	84
Qy	61	NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	120
Db	85	NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV	144
Qy	121	DVNCCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAACLLP	180
Db	145	DVNCCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAACLLP	204
Qy	181	KLDELDEGKASAKQRLKASLOKGERAFKAWARLSORFPKAEFAEVS KLVDLT	240
Db	205	KLDELDEGKASAKQRLKASLOKGERAFKAWARLSORFPKAEFAEVS KLVDLT	264
Qy	241	VHTECCHGDLLECADRADLAKYICENODSISSKLKECCBKPLEKSHCIAEVENDEMPA	300
Db	265	VHTECCHGDLLECADRADLAKYICENODSISSKLKECCBKPLEKSHCIAEVENDEMPA	324
Qy	301	DLPSLAADFVESKDVCKNVAEKDVFGLMFYIYARRHPDYSVVLRLAKTYETTLEK	360
Db	325	DLPSLAADFVESKDVCKNVAEKDVFGLMFYIYARRHPDYSVVLRLAKTYETTLEK	384
Qy	361	CAAADPHECVAKVDFEPKPLVEEPQNLIKONCELFEOLEKFKONALLVRYTKVPQVST	420
Db	385	CAAADPHECVAKVDFEPKPLVEEPQNLIKONCELFEOLEKFKONALLVRYTKVPQVST	444
Qy	421	PTLVEVSRNKGKSGCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES	480
Db	445	PTLVEVSRNKGKSGCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES	504
Qy	481	LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVKHKPKAT	540
Db	505	LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVKHKPKAT	564
Qy	541	KEQLKAVMDFFAFAVEKCCKADKCTCFABEGKLVAAASQAALGLMSPRLEVP	593
Db	565	KEQLKAVMDFFAFAVEKCCKADKCTCFABEGKLVAAASQAALGLVA---ETPTPWRDA	621
Qy	594	-----CSHALPQGLSPGQVIIVRGLVLOE-----PKHFT-----	622
Db	622	ETGERLVCAQCP-----GTFVORPCRRSPPTCGPCPPRHYTFQWNYLERCR	669
Qy	623	-----VSLRDQAHA-----	632
Db	670	YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPPGAGVIAPTGPSQNTQ	729
Qy	633	-----PVTLRAS-----	639
Db	730	CQCPPTGTFSSSSSSBQCFHRNCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPGAE	789
Qy	640	-----FADRTLQWLSR-----WG-----OKKLISAPFLFYPQRF	668
Db	790	ECERAVIDFVAQDISIKRLQLLQALEAPEGPGPTPRAGRALQLKL-----RRR	840
Qy	669	FEVLLLFQEGGLKIALNGOGLGATSMNQALEQRLRLRISG 709	
Db	841	LTELGLAQDQALLVRL-----LQALRVARMFG 867	
RESULT 25			
ADFL14978			
ID	ADFL14978	standard; protein; 880 AA.	
XX	XX		
AC	ADFL14978;		
XX	XX		
DT	12-FEB-2004	(first entry)	
XX	XX		
DE	Human albumin therapeutic fusion protein SeqID274.		

albumin fusion protein; albumin activity; human serum albumin; serum osmotic pressure; shelf-life; stability; antidiabetic; gene therapy; diabetes mellitus; human.

Chimeric.

Homo sapiens.

WO2003060071-A2.

24-JUL-2003.

23-DEC-2002; 2002WO-US040891.

21-DEC-2001; 2001US-0341811P.

24-JAN-2002; 2002US-0350358P.

28-JAN-2002; 2002US-0351360P.

26-FEB-2002; 2002US-0359370P.

28-FEB-2002; 2002US-0360000P.

27-MAR-2002; 2002US-0367500P.

08-APR-2002; 2002US-0370227P.

10-MAY-2002; 2002US-0378950P.

24-MAY-2002; 2002US-0382617P.

28-MAY-2002; 2002US-0383123P.

05-JUN-2002; 2002US-0385708P.

10-JUL-2002; 2002US-0394625P.

24-JUL-2002; 2002US-0398008P.

09-AUG-2002; 2002US-0402131P.

13-AUG-2002; 2002US-0402708P.

18-SEP-2002; 2002US-0411355P.

18-SEP-2002; 2002US-0411426P.

02-OCT-2002; 2002US-0414984P.

11-OCT-2002; 2002US-0417611P.

23-OCT-2002; 2002US-0420246P.

05-NOV-2002; 2002US-0423623P.

(HUMA-) HUMAN GENOME SCI INC.

(DELZ ) DELTA BIOTECHNOLOGY LTD.

(PRIN-) PRINCIPIA PHARM CORP.

Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;

MPI; 2003-598517/56.

New albumin fusion protein, useful for preparing a composition for treating diabetes mellitus.

Example 4; SEQ ID NO 274; 24pp; English.

This invention relates to a novel albumin fusion protein having albumin or biological activity. Human serum albumin is responsible for a significant proportion of the osmotic pressure of serum and also functions as a carrier of endogenous and exogenous ligands. The fusion of albumin to a therapeutic protein may increase shelf-life and stability of the therapeutic protein. The albumin fusion protein of the invention may allow production of compositions with antidiabetic activity whilst the nucleotide sequence which encodes it may be useful for gene therapy. The albumin fusion protein is useful for preparing a composition for treating diabetes mellitus. The present sequence is the amino acid sequence of a novel full-length human albumin therapeutic fusion protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/publishedpct\\_sequences](http://ftp.wipo.int/pub/publishedpct_sequences)

Sequence 880 AA;

Query Match 83.2%; Score 3151; DB 7; Length 880;

Best Local Similarity 70.5%; Pred. No. 8.8e-73;

Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHKSEVAHRFKDGLGENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADES 60

Db 25 DAHKSEVAHRFKDGLGENFKALVLIIFAQYLOQCPFDHVKLVNEVTEFAKTCVADES 84

Qy	61	NCDSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV	120
Db	85	NCDSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV	144
Qy	121	DVMTAFHNEBETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP	180
Db	145	DVMTAFHNEBETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP	204
Qy	181	KLDELDEGKASSAKQRLKASLQKFGERAFAKAWARLSORPPKAFPAEVSKLVTDLTK	240
Db	205	KLDELDEGKASSAKQRLKASLQKFGERAFAKAWARLSORPPKAFPAEVSKLVTDLTK	264
Qy	241	VHTECCGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDMPA	300
Db	265	VHTECCGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDMPA	324
Qy	301	DLPSLAADFVSKDVKCNKABAKOVFLGMFLYIYARRHPDYVSVLLRLAKTYETTLK	360
Db	325	DLPSLAADFVSKDVKCNKABAKOVFLGMFLYIYARRHPDYVSVLLRLAKTYETTLK	384
Qy	361	CAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPOVST	420
Db	385	CAADPHCEYAKVDFEFPKPLVEEPQNLIKQNCFLPEQLGEYKFNALLVRYTKVPOVST	444
Qy	421	PTLVEVSRLKGVSKCCCKHPKAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCTTES	480
Db	445	PTLVEVSRLKGVSKCCCKHPKAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCTTES	504
Qy	481	LVNRRPCFSALEVDYVYKFNATFTFFHADICTLSEKQIKQKQALVELVGHKPKAT	540
Db	505	LVNRRPCFSALEVDYVYKFNATFTFFHADICTLSEKQIKQKQALVELVGHKPKAT	564
Qy	541	KEQLKAVMDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGMSPLRLEVP	593
Db	565	KEQLKAVMDFAAFVEKCKADDDKTCFAEKGKLVAAASQAALGMSPLRLEVP	621
Qy	594	-----C\$HALPQGLSPQVILVRLQLE-----PKHFT-----	622
Db	622	ETGERLVCAQCFP-----GTFVQPCRRDSTTCGCPPRHYTQFMVLERCR	669
Qy	623	-----VSLRDQAANA-----	632
Db	670	YCNVLGGEREBEARACHATHNRCRCRTGTFPAHAGFCLHASCPPGAGVIAPGTPSQNTQ	729
Qy	633	-----PVTLRAS-----	639
Db	730	CQCPPTGTFSSSSSSSQCPHRNCTALGLALNVPGSSSHDTLCTSGTFPLSTRVPGAE	789
Qy	640	-----FADRTLQWISR-----WG-----QKKLISAPFLFYQRF	668
Db	790	ECERAVIDFVAFQDISIKELQELQALPAPEGWGTPPRAGRALQLKL-----RRR	840
Qy	669	FEVLLFQEGGLKALNGQGLGATSMNQOALEQLRELISG	709
Db	841	LTELIGAQDGLLVRL-----LQALVARMPG	867
RESULT 26			
ID	ADFI4937	standard; protein; 880 AA.	
AC	ADFI4937;		
XX	12-FEB-2004	(first entry)	
DT	Human albumin therapeutic fusion protein SeqID233.		
DE	albumin fusion protein; albumin activity; human serum albumin;		
KW	serum osmotic pressure; shelf-life; stability; antidiabetic;		
KW	gene therapy; diabetes mellitus; human.		
XX	Chimeric.		
OS			
OS	Homo sapiens.		
XX	WO2003060071-A2.		
XX	24-JUL-2003.		
PD	23-DEC-2002; 2002WO-US040891.		
XX	21-DEC-2001; 2001US-0341811P.		
PR	24-JAN-2002; 2002US-0350358P.		
PR	28-JAN-2002; 2002US-0351360P.		
PR	26-FEB-2002; 2002US-0359370P.		
PR	28-FEB-2002; 2002US-0360000P.		
PR	27-MAR-2002; 2002US-0367500P.		
PR	08-APR-2002; 2002US-0370227P.		
PR	10-MAY-2002; 2002US-0378950P.		
PR	24-MAY-2002; 2002US-0382617P.		
PR	28-MAY-2002; 2002US-0383123P.		
PR	05-JUN-2002; 2002US-0385708P.		
PR	10-JUL-2002; 2002US-0394625P.		
PR	24-JUL-2002; 2002US-0398008P.		
PR	09-AUG-2002; 2002US-0402131P.		
PR	13-AUG-2002; 2002US-0402708P.		
PR	18-SEP-2002; 2002US-0411355P.		
PR	18-SEP-2002; 2002US-0411426P.		
PR	02-OCT-2002; 2002US-0414984P.		
PR	11-OCT-2002; 2002US-0417611P.		
PR	23-OCT-2002; 2002US-0420246P.		
PR	05-NOV-2002; 2002US-0423623P.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	(DELZ ) DELTA BIOTECHNOLOGY LTD.		
PA	(PRIN-) PRINCIPIA PHARM CORP.		
XX	Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;		
XX	WPI; 2003-598517/56.		
DR	New albumin fusion protein, useful for preparing a composition for		
XX	treating diabetes mellitus.		
PT	Example 4; SEQ ID NO 233; 24pp; English.		
PS	This invention relates to a novel albumin fusion protein having albumin		
XX	or biological activity. Human serum albumin is responsible for a		
CC	significant proportion of the osmotic pressure of serum and also		
CC	functions as a carrier of endogenous and exogenous ligands. The fusion of		
CC	albumin to a therapeutic protein may increase shelf-life and stability of		
CC	the therapeutic protein. The albumin fusion protein of the invention may		
CC	allow production of compositions with antidiabetic activity whilst the		
CC	nucleotide sequence which encodes it may be useful for gene therapy. The		
CC	albumin fusion protein is useful for preparing a composition for treating		
CC	diabetes mellitus. The present sequence is the amino acid sequence of a		
CC	novel full-length human albumin therapeutic fusion protein of the		
CC	invention. Note: The sequence data for this patent did not form part of		
CC	the printed specification, but was obtained in electronic format directly		
CC	from WIPO at ftp.wipo.int/pub/publishedpct_sequences		
XX	Sequence 880 AA;		
SQ	Query Match 83.2%; Score 3151; DB 7; Length 880;		
	Best Local Similarity 70.5%; Pred. No. 8.8e-73;		
	Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;		
Qy	1 DAHSEVAHREFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAG	60	
Db	25 DAHSEVAHREFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAG	84	
Qy	61 NCDSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV	120	
Db	85 NCDSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV	144	
Qy	121 DVMTAFHNEBETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP	180	

Db 145 DVNCTAFHNEETFLKKYLVEIARRHPYFYVAPPELLFFAKRYKAAFTCCQQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSQRPKFAFAVSVKLVDLTk 240  
Db 205 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSQRPKFAFAVSVKLVDLTk 264  
Qy 241 VHTTECHGDLLECADRADIAKYICENODSISKLKCECKPILLESKSHCIAEVENDEMPA 300  
Db 265 VHTTECHGDLLECADRADIAKYICENODSISKLKCECKPILLESKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYAEKXVFLGMLFYIYARRHPDYSVVLRLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYAEKXVFLGMLFYIYARRHPDYSVVLRLRLAKTYETTLK 384  
Qy 361 CAAADHECYAKVDFEPKPLVEEPONLIKQNCLEFQOLGKBFQONALLVRYTKVPQVST 420  
Db 385 CAAADHECYAKVDFEPKPLVEEPONLIKQNCLEFQOLGKBFQONALLVRYTKVPQVST 444  
Qy 421 PTILVESRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTILVESRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVKHKPKAT 540  
Db 505 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVKHKPKAT 564  
Qy 541 KEOLKAVMDFFAFAVEKCKKADKCTCFABEGKLVAAASQAALGLMSPRLEVP----- 593  
Db 565 KEOLKAVMDFFAFAVEKCKKADKCTCFABEGKLVAAASQAALGLVA---ETPTYPWRDA 621  
Qy 594 -----CSHALPQGLSPGVIIVRGLVLOE-----PKHET----- 622  
Db 622 ETGERLVCAQCP-----GTFVQPCRRDSPTTCGCPKPHYTFQWNYLERCR 669  
Qy 623 -----VSLRDQAAHA----- 632  
Db 670 YCNVLGGEREEARACHATHNACRCHTGFFAHAGFCLEHASCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVLTRAS----- 639  
Db 730 CQPCPPGTFSSASSSSSQCPHRQCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPGA 789  
Qy 640 -----FADRTLQWISR-----WG-----OKKLISAPFLFYPORF 668  
Db 790 ECERAVLDFVAFQDISIRLQRLLOALEAPEGKGPTRAGRAALQLKL-----RRR 840  
Qy 669 FEVLLLFQEGGLKIALNGOGLGATSMNQALQLELRLISG 709  
Db 841 LTELGLAQDQALLVRL-----LQALRVARMPG 867

RESULT 27  
ADFL14921  
ID ADF14921 standard; protein; 880 AA.  
XX  
AC ADF14921;  
XX  
XX 12-FEB-2004 (first entry)  
XX Human albumin therapeutic fusion protein SeqID217.  
XX  
XX albumin fusion protein; albumin activity; human serum albumin;  
XX serum osmotic pressure; shelf-life; stability; antidiabetic;  
XX gene therapy; diabetes mellitus; human.  
XX Chimeric.  
OS Homo sapiens.  
XX  
XX W02003060071-A2.  
PN  
XX  
PD 24-JUL-2003.  
XX

PF 23-DEC-2002; 2002WO-US040891.  
XX  
PR 21-DEC-2001; 2001US-0341811P.  
PR 24-JAN-2002; 2002US-0350358P.  
PR 26-JAN-2002; 2002US-0351360P.  
PR 28-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.  
PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-MAY-2002; 2002US-0382617P.  
PR 28-MAY-2002; 2002US-0383123P.  
PR 05-JUN-2002; 2002US-0385708P.  
PR 10-JUL-2002; 2002US-0394625P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 18-SEP-2002; 2002US-0411426P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX (PRIN-) PRINCIPIA PHARM CORP.  
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX  
XX New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX  
XX Example 4; SEQ ID NO 217; 24pp; English.  
XX  
XX This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
XX significant proportion of the osmotic pressure of serum and also  
XX functions as a carrier of endogenous and exogenous ligands. The fusion of  
XX albumin to a therapeutic protein may increase shelf-life and stability of  
XX the therapeutic protein. The albumin fusion protein of the invention may  
XX allow production of compositions with antidiabetic activity whilst the  
XX nucleotide sequence which encodes it may be useful for gene therapy. The  
XX albumin fusion protein is useful for preparing a composition for treating  
XX diabetes mellitus. The present sequence is the amino acid sequence of a  
XX novel full-length human albumin therapeutic fusion protein of the  
XX invention. Note: The sequence data for this patent did not form part of  
XX the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/publishedpct\_sequences  
XX  
XX Sequence 880 AA;  
XX  
XX Query Match 83.2%; Score 3151; DB 7; Length 880;  
XX Best Local Similarity 70.5%; Pred. No. 8.9e-73;  
XX Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHKSEVAHRFKD LGENFKALVLI AFAQYLOQCPFDHVKLVNVE TFAKTCVADES AE 60  
Db 25 DAHKSEVAHRFKD LGENFKALVLI AFAQYLOQCPFDHVKLVNVE TFAKTCVADES AE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGMDCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGMDCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
Qy 121 DVNCTAFHNEETFLKKYLVEIARRHPYFYVAPPELLFFAKRYKAAFTCCQQAADKAACLLP 180  
Db 145 DVNCTAFHNEETFLKKYLVEIARRHPYFYVAPPELLFFAKRYKAAFTCCQQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSQRPKFAFAVSVKLVDLTk 240  
Db 205 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSQRPKFAFAVSVKLVDLTk 264

```
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCECKEPLLEKSHCIAEVENDMPA 300
DB 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCECKEPLLEKSHCIAEVENDMPA 324
QY 301 DLPSLAADFVSKDVCKNRYAKADVFLGMFLYEVARRHPDYVSVLLRLAKTYETILLEK 360
DB 325 DLPSLAADFVSKDVCKNRYAKADVFLGMFLYEVARRHPDYVSVLLRLAKTYETILLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPPQNLKONCELFEOQGEYKFNALLVRYTKKVPQVST 420
DB 385 CAADPHCEYAKVDFEFPKPLVEEPPQNLKONCELFEOQGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSNNLKVSGKCKCHPEAKRMPCAEDYLSVNLQNLVLEHKTVPVSDRVTKCTTES 480
DB 445 PTLVEVSNNLKVSGKCKCHPEAKRMPCAEDYLSVNLQNLVLEHKTVPVSDRVTKCTTES 504
QY 481 LVNRRPCFSALEVDYVVPKFNFAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDYVVPKFNFAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAERGGKLVAAASQAALGLMSPLRLEVP----- 593
DB 565 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAERGGKLVAAASQAALGLVA---ETPTYPWRDA 621
QY 594 -----CSHALPQGLSPGVIIVRGLVLQSE-----PKHFT----- 622
DB 622 ETGERLVCAQCP-----GTFFVQPCRRDSTTCGCPCPRHRYTQFVWYLERCR 669
QY 623 -----VSLRDOAAHA----- 632
DB 670 YCNVLGEBEAREACHATHNACRCRTGFFAHAGFCLEHASCPPGAGVIAPGTPSQNTQ 729
QY 633 -----PVTLRAS----- 639
DB 730 CQPCPPGTFSSSSSEQCPHRNCTALGLALNVGSSSHDTLCTSGTFPLSTRVPAE 789
QY 640 -----FADRTLOWISR-----WG-----QKLIASIFLFPQRF 668
DB 790 ECERAVIDFVAFQDISIKRLQELLOALAPEGWGTPPRAGRAALQKL-----RRR 840
QY 669 FEVLLFQEGGLKALNGQGLGATSMNQOALEQLRELISG 709
DB 841 LTELIGAQDGLLVRL-----LQALRVARMGP 867

RESULT 28
ADH21282
ID ADH21282 standard; protein; 880 AA.
XX AC ADH21282;
XX DT 11-MAR-2004 (first entry)
XX DE Human albumin/mutant TR6 fusion protein, SEQ ID NO:79.
XX KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy.
XX OS Synthetic.
OS Chimeric.
OS Homo sapiens.
XX PN WO2003059934-A2.
XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002WO-US040892.
```

```
XX 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0414984P.
PR 02-OCT-2002; 2002US-0417611P.
PR 11-OCT-2002; 2002US-0420246P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Haseltine WA;
XX WPI; 2003-598501/56.
XX New albumin fusion protein, useful for preparing a composition for
XX treating diabetes mellitus.
XX Disclosure; SEQ ID NO 79; 1086pp; English.
XX The invention relates to fusion proteins comprising human serum albumin
XX (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
XX antibody or peptide or their variants or fragments. The therapeutic
XX protein may be fused to the N-terminus, the C-terminus or both termini of
XX albumin via a linker. The albumin component of the fusion proteins
XX prolongs the shelf-life and the in vitro and vivo biological activity of
XX the proteins compared with those of the corresponding therapeutic
XX proteins on their own. The invention also relates to nucleic acids
XX encoding albumin fusion proteins, vectors and host cells comprising an
XX albumin fusion protein nucleic acid, compositions and kits comprising an
XX albumin fusion protein, the method of extending the shelf-life of a
XX therapeutic protein by fusion with albumin, and the treatment of disease
XX using an albumin fusion protein. The albumin fusion proteins may be used
XX in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
XX related conditions. Specifically the albumin fusion proteins may be used
XX to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
XX (especially neuropathy), retinopathy, cardiovascular disorders
XX (especially heart disease, renal disorders and obesity). The proteins may
XX also be used in a method of maintaining a basal glucose level in a
XX patient and in a method for losing weight. The present sequence is
XX related to the invention.
XX SQ Sequence 880 AA;
XX Query Match 83.2%; Score 3151; DB 7; Length 880;
XX Best Local Similarity 70.5%; Pred. No. 8.8e-73;
XX Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;
QY 1 DAHSEVAHRFKDLGEENFKALVLIATAFYQLQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIATAFYQLQCPFFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREPV 120
DB 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREPV 144
QY 121 DVMTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFPTECCQADKAACLLP 180
DB 145 DVMTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFPTECCQADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSQRFPAEFAEVSCLVDTLTK 240
DB 205 KLDELDEGKASSAKQRLKCSAQKFGGERAFKAWAVARLSQRFPAEFAEVSCLVDTLTK 264
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKCECKEPLLEKSHCIAEVENDMPA 300
```

```
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFAYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFAYARRHPDYSVLLRLAKTYETTTLEKC 384
Qy 361 CAAADPHCEYAKVDFEPKPIVERPQNLIKONCELFEOLGKYKFONALLVRYTKVPQVST 420
Db 385 CAAADPHCEYAKVDFEPKPIVERPQNLIKONCELFEOLGKYKFONALLVRYTKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDVETVVPKFEFNAETFTFHADICTLSEKERQIKQTALVELVHKRPKAT 540
Db 505 LVNRRPCFSALEVDVETVVPKFEFNAETFTFHADICTLSEKERQIKQTALVELVHKRPKAT 564
Qy 541 KEOLKAVMDDFAAFVEKCKKADDKETCTFAEEGKLVAAASQAALGLMSPRLEVP----- 593
Db 565 KEOLKAVMDDFAAFVEKCKKADDKETCTFAEEGKLVAAASQAALGLVA---ETPTYWRDA 621
Qy 594 -----CSHALPQGLSPGQVIVRGLVLOE-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVQPCRRDSTPTTCGCPRHYTQFWNYLERCR 669
Qy 623 -----VSLRQAAHA----- 632
Db 670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFLEHASCPGAGVIAPGTPSQNTQ 729
Qy 633 -----PVTLRAS----- 639
Db 730 CQCPPTGFSASSSSBQCPHQCTALGLALNVPGSSSHDTLCTSGTGPPLSTRVPGAE 789
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYQRF 668
Db 790 ECERAVIDFVAFQDISIKRLQLQALEPFGWGPTPRAGRAALQLKL-----RRR 840
Qy 669 FEVILLFOEGGLKIALNGQGLGATSMNQALEOLRELRISG 709
Db 841 LTELGAQDQALLVRL-----LQALRVARMFG 867

RESULT 29
ADH21273
ID ADH21273 standard; protein; 880 AA.
XX
AC ADH21273;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human albumin/TR6 fragment fusion protein, SEQ ID NO:70.
XX
KW Fusion protein; human serum albumin; HSA; therapeutic protein;
KW shelf-life; in vitro biological activity; in vivo biological activity;
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
KW anorectic; ophthalmological; gene therapy.
XX
OS Chimeric.
OS Homo sapiens.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
PR 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
```

```
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
XX WPI; 2003-598501/56.
XX
DR New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 70; 1086pp; English.
XX
CC The invention relates to fusion proteins comprising human serum albumin
CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
CC antibody or peptide or their variants or fragments. The therapeutic
CC protein may be fused to the N-terminus, the C-terminus or both termini of
CC albumin via a linker. The albumin component of the fusion proteins
CC prolongs the shelf-life and the in vitro and vivo biological activity of
CC the proteins compared with those of the corresponding therapeutic
CC proteins on their own. The invention also relates to nucleic acids
CC encoding albumin fusion proteins, vectors and host cells comprising an
CC albumin fusion protein nucleic acid, compositions and kits comprising an
CC albumin fusion protein, the method of extending the shelf-life of a
CC therapeutic protein by fusion with albumin, and the treatment of disease
CC using an albumin fusion protein. The albumin fusion proteins may be used
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
CC related conditions. Specifically the albumin fusion proteins may be used
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
CC (especially neuropathy), retinopathy, cardiovascular disorders
CC (especially heart disease, renal disorders and obesity. The proteins may
CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.
XX
SQ Sequence 880 AA;
```

```
Query Match 83.2%; Score 3151; DB 7; Length 880;
Best Local Similarity 70.5%; Pred. No. 8.8e-73;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRFKDLGGEENFKALVLIATAFYLOQCPPEHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGGEENFKALVLIATAFYLOQCPPEHVKLVNEVTEFAKTCVADESAE 84
Qy 61 NDKSLHTLFGKLCITVATIRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
Db 85 NDKSLHTLFGKLCITVATIRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKLYETARRHPYFVAPPELLFFAKRYKAAFTCCQADKAACLIP 180
Db 145 DVMCTAFHDNEETFLKKLYETARRHPYFVAPPELLFFAKRYKAAFTCCQADKAACLIP 204
Qy 181 KLDELDEGKASSAKQRLKASIQKFGERAFAKAWAVARLSQRFPFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKASIQKFGERAFAKAWAVARLSQRFPFAEVSCLVTDLT 264
Qy 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVENDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVENDEMPA 324
```



Qy 301 DLPSLAADFVESKDVCKNYABAKDVLGMLFYIYARRHPDYSVVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYABAKDVLGMLFYIYARRHPDYSVVLLRLAKTYETTLK 384  
Qy 361 CAADPHCYAKVDEKPLVEEPPONLTKONCELFQELGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHCYAKVDEKPLVEEPPONLTKONCELFQELGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETFTTHADICTLSEKEROIKQOTALVELVHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYKPEFNAETFTTHADICTLSEKEROIKQOTALVELVHKPKAT 564  
Qy 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAQAALGLMSPRLEVP----- 593  
Db 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAQAALGLVA---ETPTYWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIVRGLVLQE-----PKHFT----- 622  
Db 622 ETGERLCAQCP-----GTFVQPCRRDSEPTTCGPCPRHYTQFWNLYRCR 669  
Qy 623 -----VSLRDOAHV----- 632  
Db 670 YCNVLGCBEBEARACHATHNRCRCRTGFFAHAGFCLHASCPCPGAGVAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 CQPCPPGTFSSSSSSQCPHRNCTALGLALNVGSSSHDTLCTSGFPPLSTRVGA 789  
Qy 640 -----FADRTLOWISR-----WG-----QKKLISAPFLFYQRF 668  
Db 790 ECERAVIDFVAFQDISIKRLQLLQALRAPEGWGTTPAGRAALQLKL-----RRR 840  
Qy 669 FEVLLFQGGKIALNGOGILGATSMNQOALEQLRELISG 709  
Db 841 LTELGAQDGLLVRL-----LQALRVARMPG 867

RESULT 30  
ADH21293  
ID ADH21293 standard; protein; 880 AA.  
XX ADH21293;  
AC ADH21293;  
XX 11-MAR-2004 (first entry)  
DT Human albumin/TR6 fragment fusion protein, SEQ ID NO:90.  
XX Fusion protein; human serum albumin; HSA; therapeutic protein;  
XX shelf-life; in vitro biological activity; in vivo biological activity;  
KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;  
KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;  
KW retinopathy; cardiovascular disorder; heart disease; renal disorder;  
KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiac;  
KW anorectic; ophthalmological; gene therapy.  
XX Chimeric.  
OS Homo sapiens.  
PN WO2003059934-A2.  
XX 24-JUL-2003.  
PD 23-DEC-2002; 2002WO-US040892.  
XX 21-DEC-2001; 2001US-0341811P.  
XX 24-JAN-2002; 2002US-0350358P.  
PR 26-FEB-2002; 2002US-0359370P.  
PR 28-FEB-2002; 2002US-0360000P.  
PR 27-MAR-2002; 2002US-0367500P.

PR 08-APR-2002; 2002US-0370227P.  
PR 10-MAY-2002; 2002US-0378950P.  
PR 24-JUL-2002; 2002US-0398008P.  
PR 09-AUG-2002; 2002US-0402131P.  
PR 13-AUG-2002; 2002US-0402708P.  
PR 18-SEP-2002; 2002US-0411355P.  
PR 02-OCT-2002; 2002US-0414984P.  
PR 11-OCT-2002; 2002US-0417611P.  
PR 23-OCT-2002; 2002US-0420246P.  
PR 05-NOV-2002; 2002US-0423623P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Haseltine WA;  
PI WPI; 2003-598501/56.  
XX New albumin fusion protein, useful for preparing a composition for  
PT treating diabetes mellitus.  
XX Disclosure; SEQ ID NO 90; 1086pp; English.  
XX The invention relates to fusion proteins comprising human serum albumin  
XX (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,  
CC antibody or peptide or their variants or fragments. The therapeutic  
CC protein may be fused to the N-terminus, the C-terminus or both termini of  
CC albumin via a linker. The albumin component of the fusion proteins  
CC prolongs the shelf-life and the in vitro and vivo biological activity of  
CC the proteins compared with those of the corresponding therapeutic  
CC proteins on their own. The invention also relates to nucleic acids  
CC encoding albumin fusion proteins, vectors and host cells comprising an  
CC albumin fusion protein nucleic acid, compositions and kits comprising an  
CC albumin fusion protein, the method of extending the shelf-life of a  
CC therapeutic protein by fusion with albumin, and the treatment of disease  
CC using an albumin fusion protein. The albumin fusion proteins may be used  
CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-  
CC related conditions. Specifically the albumin fusion proteins may be used  
CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders  
CC (especially neuropathy), retinopathy, cardiovascular disorders  
CC (especially heart disease), renal disorders and obesity. The proteins may  
CC also be used in a method of maintaining a basal glucose level in a  
CC patient and in a method for losing weight. The present sequence is  
CC related to the invention.  
XX Sequence 880 AA;  
SQ Query Match 83.2%; Score 3151; DB 7; Length 880;  
Best Local Similarity 70.5%; Pred. No. 8.8e-73;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHSEVAHRFKDILGEENFKALVLIAPAFQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60  
Db 25 DAHSEVAHRFKDILGEENFKALVLIAPAFQYLQCCPFEDHVKLVNEVTEFAKTCVADES 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETTGEMADCCAKOEPERNECFLOHKDNNPLRLVREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETTGEMADCCAKOEPERNECFLOHKDNNPLRLVREV 144  
Qy 121 DVMCTAFHDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQADRAACLLP 180  
Db 145 DVMCTAFHDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQADRAACLLP 204  
Qy 181 KLDELDEGKASSAKQRLKCAKSLQKFGERAFAKAWAVARLSQRFPAEFAVSKLVTDLTK 240  
Db 205 KLDELDEGKASSAKQRLKCAKSLQKFGERAFAKAWAVARLSQRFPAEFAVSKLVTDLTK 264  
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLCCKCEKPILEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLCCKCEKPILEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYABAKDVLGMLFYIYARRHPDYSVVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYABAKDVLGMLFYIYARRHPDYSVVLLRLAKTYETTLK 384

```

QY 361 CAADPHECYAKVDFEPKPLAVERPQNLIKONCELFQOLGEYKFEONALLVRYTKKVPQVST 420
Db 385 CAADPHECYAKVDFEPKPLAVERPQNLIKONCELFQOLGEYKFEONALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTCKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTCKCTES 504
QY 481 LVNRRPCFSALEVDETVVPKFNATFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETVVPKFNATFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFEVKCCCKADDDKCTCFABEGKKLVAAQAALGLMSPRLEVP----- 593
Db 565 KEQLKAVMDDFAAFEVKCCCKADDDKCTCFABEGKKLVAAQAALGLVA---ETPTYPMRDA 621
QY 594 -----CSHALPQGLSPGQVIIVRGLVLQ-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVORPCERDSPTTCGPPRHHYTOFWNYLERCR 669
QY 623 -----VSLRDOAAHA----- 632
Db 670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLHASCPPGAGVIAPGTPSQNTQ 729
QY 633 -----PVTLEAS----- 639
Db 730 QPCPPGTFSSSSSSQCPHRNCTALGLALNVGSSSHDTLCTSTGTPPLSTRVPGAE 789
QY 640 -----FADRTLQWISR-----WG-----QKKLISAPELFYPPQRF 668
Db 790 ECERAVIDFVAFQDISIKRLQRLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
QY 669 FEVILLFQEGGLKALNGQIGATSMNQALQOLRELIRSG 709
Db 841 LTELGAQDGALLVRL-----LQALRVARMFG 867

```

Search completed: October 17, 2005, 08:29:01  
 Job time : 180 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 43 Seconds  
(without alignments)  
1246.466 Million cell updates/sec

Title: US-10-933-523-18  
Perfect score: 3785  
Sequence: 1 DAHKEVAHRPKDLGEENFK.....leglreirsgvqlcvs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3140.6	83.0	783	1	US-08-256-938-2
2	3114.9	82.3	978	4	US-08-897-956A-3
3	3103	82.0	585	1	US-08-153-799-14
4	3103	82.0	585	2	US-08-702-572-2
5	3103	82.0	585	3	US-08-769-746-2
6	3103	82.0	609	4	US-09-976-594-977
7	3103	82.0	609	4	US-09-919-039-370
8	3103	82.0	610	2	US-08-797-689-2
9	3103	82.0	610	4	US-09-984-186-2
10	3103	82.0	622	4	US-09-949-016-11170
11	3103	82.0	787	1	US-08-256-938-4
12	3103	82.0	787	2	US-08-797-689-16
13	3103	82.0	787	4	US-09-984-186-16
14	3099	81.9	609	1	US-08-222-619-3
15	3099	81.9	609	1	US-08-433-037-4
16	3099	81.9	609	4	US-08-897-956A-2
17	3099	81.9	609	5	PCT-US95-04075-3
18	3093	81.7	585	1	US-08-448-196A-3
19	3093	81.7	585	2	US-08-984-176-1
20	2458.9	65.0	583	1	US-08-448-196A-5
21	2450.9	64.8	583	1	US-08-448-196A-4
22	2432.9	64.3	583	1	US-08-448-196A-6
23	2426	64.1	584	1	US-08-448-196A-7
24	2389.8	63.1	582	1	US-08-134-638-1
25	1252.7	33.1	609	1	US-08-222-619-5
26	1252.7	33.1	609	4	US-09-976-594-456
27	1252.7	33.1	609	5	PCT-US95-04075-4

28	1252.7	33.1	612	4	US-09-949-016-11201	Sequence 11201, A
29	1211.3	32.0	609	4	US-09-186-949A-2	Sequence 2, Appli
30	1210.9	32.0	590	2	US-08-377-309-2	Sequence 2, Appli
31	1210.9	32.0	590	3	US-09-186-723-2	Sequence 5, Appli
32	1210.9	32.0	590	3	US-08-505-012-5	Sequence 5, Appli
33	1210.9	32.0	590	4	US-09-186-949A-3	Sequence 3, Appli
34	1210.9	32.0	590	4	US-08-758-757-2	Sequence 2, Appli
35	1210.9	32.0	590	4	US-09-187-978-2	Sequence 2, Appli
36	1210.9	32.0	590	4	US-10-115-701A-2	Sequence 2, Appli
37	1210.9	32.0	590	4	US-09-940-308A-2	Sequence 2, Appli
38	1210.9	32.0	590	4	US-09-940-308A-2	Sequence 2, Appli
39	1210.9	32.0	590	5	PCT-US96-00996-5	Sequence 5, Appli
40	1168.1	30.9	579	1	US-08-448-196A-8	Sequence 8, Appli
41	1059	28.0	599	1	US-08-222-619-2	Sequence 2, Appli
42	1059	28.0	599	3	US-08-221-767-24	Sequence 24, Appli
43	1059	28.0	599	5	PCT-US95-04075-2	Sequence 2, Appli
44	926	24.5	393	2	US-08-377-309-7	Sequence 7, Appli
45	926	24.5	393	3	US-09-186-723-7	Sequence 7, Appli
46	926	24.5	393	3	US-08-505-012-10	Sequence 10, Appli
47	926	24.5	393	4	US-09-186-949A-8	Sequence 8, Appli
48	926	24.5	393	4	US-08-758-757-7	Sequence 7, Appli
49	926	24.5	393	4	US-09-187-978-7	Sequence 7, Appli
50	926	24.5	393	4	US-10-115-701A-7	Sequence 7, Appli
51	926	24.5	393	4	US-09-940-308A-7	Sequence 7, Appli
52	926	24.5	393	4	US-09-940-308A-7	Sequence 7, Appli
53	926	24.5	393	5	PCT-US96-00996-10	Sequence 10, Appli
54	777	20.5	324	3	US-08-505-012-12	Sequence 12, Appli
55	777	20.5	324	5	PCT-US96-00996-12	Sequence 12, Appli
56	777	20.5	325	2	US-08-377-309-8	Sequence 8, Appli
57	777	20.5	325	3	US-09-186-723-8	Sequence 8, Appli
58	777	20.5	325	3	US-08-505-012-11	Sequence 11, Appli
59	777	20.5	325	4	US-09-186-949A-9	Sequence 9, Appli
60	777	20.5	325	4	US-08-758-757-8	Sequence 8, Appli
61	777	20.5	325	4	US-09-187-978-8	Sequence 8, Appli
62	777	20.5	325	4	US-10-115-701A-8	Sequence 8, Appli
63	777	20.5	325	4	US-09-940-308A-8	Sequence 8, Appli
64	777	20.5	325	4	US-09-940-308A-8	Sequence 8, Appli
65	777	20.5	325	5	PCT-US96-00996-11	Sequence 11, Appli
66	765.3	20.2	590	1	US-08-448-196A-9	Sequence 9, Appli
67	709.5	18.7	336	3	US-09-131-648-1	Sequence 1, Appli
68	708.3	18.7	275	3	US-09-557-170A-25	Sequence 25, Appli
69	704.9	18.6	286	4	US-09-557-170A-27	Sequence 27, Appli
70	688.9	18.2	389	2	US-08-377-309-6	Sequence 6, Appli
71	688.9	18.2	389	3	US-09-186-723-6	Sequence 6, Appli
72	688.9	18.2	389	3	US-08-505-012-9	Sequence 9, Appli
73	688.9	18.2	389	4	US-09-186-949A-7	Sequence 7, Appli
74	688.9	18.2	389	4	US-08-758-757-6	Sequence 6, Appli
75	688.9	18.2	389	4	US-09-187-978-6	Sequence 6, Appli
76	688.9	18.2	389	4	US-10-115-701A-6	Sequence 6, Appli
77	688.9	18.2	389	4	US-09-940-308A-6	Sequence 6, Appli
78	688.9	18.2	389	4	US-09-940-308A-6	Sequence 6, Appli
79	688.9	18.2	389	5	PCT-US96-00996-9	Sequence 9, Appli
80	676	17.9	133	4	US-09-557-170A-2	Sequence 2, Appli
81	523	13.8	201	2	US-08-377-309-5	Sequence 5, Appli
82	523	13.8	201	3	US-09-186-723-5	Sequence 5, Appli
83	523	13.8	201	3	US-08-505-012-8	Sequence 8, Appli
84	523	13.8	201	4	US-09-186-949A-6	Sequence 6, Appli
85	523	13.8	201	4	US-08-758-757-5	Sequence 5, Appli
86	523	13.8	201	4	US-09-187-978-5	Sequence 5, Appli
87	523	13.8	201	4	US-10-115-701A-5	Sequence 5, Appli
88	523	13.8	201	4	US-09-940-308A-5	Sequence 5, Appli
89	523	13.8	201	4	US-09-940-308A-5	Sequence 5, Appli
90	523	13.8	201	5	PCT-US96-00996-8	Sequence 8, Appli
91	521	13.8	117	1	US-08-448-196A-2	Sequence 2, Appli
92	516	13.6	109	1	US-08-448-196A-1	Sequence 1, Appli
93	421	11.1	474	4	US-09-949-016-11630	Sequence 11630, A
94	417.5	11.0	458	6	5177002-2	Patent No. 5177002
95	417.5	11.0	458	6	5177002-2	Patent No. 5177002
96	417.5	11.0	474	1	US-08-222-619-5	Sequence 5, Appli
97	417.5	11.0	474	5	PCT-US95-04075-5	Sequence 5, Appli
98	416	11.0	458	4	US-08-618-485B-1	Sequence 1, Appli
99	416	11.0	458	6	5177002-1	Patent No. 5177002
100	416	11.0	458	6	5177002-1	Patent No. 5177002

ALIGNMENTS

RESULT 1  
US-08-256-938-2  
; Sequence 2, Application US/08256938  
; Patent No. 566863  
; GENERAL INFORMATION:  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE  
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.0 (Patentin)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,938  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92/01065  
; FILING DATE: 31-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goodman, Rosanne  
; REGISTRATION NUMBER: 32,534  
; REFERENCE/DOCKET NUMBER: ST92007-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3817  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 783 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-256-938-2

Query Match 83.0%; Score 3140.6; DB 1; Length 783;  
Best Local Similarity 80.0%; Pred. No. 5e-85;  
Matches 617; Conservative 13; Mismatches 27; Indels 114; Gaps 9;  
  
QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAOYLOQCPPEHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHRFKDGLGENFKALVLIAPAOYLOQCPPEHVKLVNEVTEFAKTCVADESAE 84  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
  
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 204  
  
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 264  
  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVVLLRLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVVLLRLAKTYETTTLEKC 384  
QY 361 CAAADPHECYAKVDFDEFKPLVEEPQNLIKQNCBFLFQQLGEYKFNQALLVRYTKKVPQVST 420  
DB 385 CAAADPHECYAKVDFDEFKPLVEEPQNLIKQNCBFLFQQLGEYKFNQALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVVNLQNLVULHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCCCKHPEAKRMPCAEDYLSVVNLQNLVULHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPEKFNAEFTTHADICTLSEKEROIKKOTALVELVKKHPKAT 540  
DB 505 LVNRRPCFSALEVDETYVPEKFNAEFTTHADICTLSEKEROIKKOTALVELVKKHPKAT 564  
QY 541 KEQLKAVMDDFAAFVEKCKCADDKETCFABEGKGLVAASQAALGL--MSFRLVPPCSHAL 598  
DB 565 KEQLKAVMDDFAAFVEKCKCADDKETCFABEGKGLVAASQAALGLTPLGPASSLPQSFL 624  
QY 599 P-----QGLSPGVIIIVRGLVLOEPKHFTVSLRDQAHAHAPVTLRASFA DRTLQWISR 650  
DB 625 KCLEQVRKIQG-----DGAALQOE----- 642  
QY 651 WGQKKLISAPFLFYQRPFFEVLLLFQEGGL-----KLALNG-----Q 687  
DB 643 ----KLCATYKLCHEP---ELVLLHSLGIPWAPLSSCFPSQALQAGCLSLQHSGLFLYQ 695  
QY 688 GLGATSMNQALE-----QRLRELIRSGSVQ 712  
DB 696 GL-----LQALEGISPELGPTLDTQLDVADPATTIWOQMEELGWAPALQ 740

RESULT 2  
US-08-897-956A-3  
; Sequence 3, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/GPA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; CURRENT FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 978  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Fusion polypeptide  
US-08-897-956A-3

Query Match 82.3%; Score 3114.9; DB 4; Length 978;  
Best Local Similarity 83.5%; Pred. No. 4.2e-84;  
Matches 608; Conservative 9; Mismatches 30; Indels 81; Gaps 8;  
  
QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAOYLOQCPPEHVKLVNEVTEFAKTCVADESAE 60  
DB 212 DAHSEVAHRFKDGLGENFKALVLIAPAOYLOQCPPEHVKLVNEVTEFAKTCVADESAE 271  
  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 331  
  
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180  
DB 332 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 391

181 KLDLDEGRKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFKAEFAVSKLVTDLTK 240  
182 KLDLDEGRKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFKAEFAVSKLVTDLTK 240  
392 KLDLDEGRKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFKAEFAVSKLVTDLTK 451  
393 KLDLDEGRKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFKAEFAVSKLVTDLTK 451  
241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVENDEMPA 300  
242 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVENDEMPA 300  
452 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVENDEMPA 511  
453 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVENDEMPA 511  
301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYISVLLRLAKTYETTLK 360  
302 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYISVLLRLAKTYETTLK 360  
512 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYISVLLRLAKTYETTLK 571  
513 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYISVLLRLAKTYETTLK 571  
361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQGEYKFNALLVRYTKVPQVST 420  
362 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQGEYKFNALLVRYTKVPQVST 420  
572 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQGEYKFNALLVRYTKVPQVST 631  
573 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQGEYKFNALLVRYTKVPQVST 631  
421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLEHKEKTPVSDRVTKCCTES 480  
422 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLEHKEKTPVSDRVTKCCTES 480  
632 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLEHKEKTPVSDRVTKCCTES 691  
633 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLEHKEKTPVSDRVTKCCTES 691  
481 LVNRRPCFSALVEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVKHKPKAT 540  
482 LVNRRPCFSALVEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVKHKPKAT 540  
692 LVNRRPCFSALVEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVKHKPKAT 751  
693 LVNRRPCFSALVEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVKHKPKAT 751  
541 KEQLKAVMDDFAAFVEKCKADDKETCFPAEKGKLVAAQALGMSPLRLEVPCHALPQ 600  
542 KEQLKAVMDDFAAFVEKCKADDKETCFPAEKGKLVAAQALGMSPLRLEVPCHALPQ 600  
752 KEQLKAVMDDFAAFVEKCKADDKETCFPAEKGKLVAAQALGMSPLRLEVPCHALPQ 795  
753 KEQLKAVMDDFAAFVEKCKADDKETCFPAEKGKLVAAQALGMSPLRLEVPCHALPQ 795  
601 GLSPGOVIIVGLVLOEPKHFVTSIRDOAAHAP-----VTLRASPADRTLQWISRW 651  
602 GLSPGOVIIVGLVLOEPKHFVTSIRDOAAHAP-----VTLRASPADRTLQWISRW 651  
796 -----GGGSPVQPKP-----VSL-----NPPWNRIFKGENVTLCN----- 826  
797 -----GGGSPVQPKP-----VSL-----NPPWNRIFKGENVTLCN----- 826  
652 GQKQLISAPPLFYFQRPPEV--LLLFQBGGLKALNGQGLGATSMNQALPOLRELRTSG 709  
653 GQKQLISAPPLFYFQRPPEV--LLLFQBGGLKALNGQGLGATSMNQALPOLRELRTSG 709  
827 -----GNNFPEVSTKWFHNGSLSEETN-----SSLN-----IVNAKPED 861  
828 -----GNNFPEVSTKWFHNGSLSEETN-----SSLN-----IVNAKPED 861  
710 SVQLYCVH 717  
862 SGEYKQ 869

RESULT 3  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB90/00650  
FILING DATE: 26-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/775952  
FILING DATE: 29-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Swope, R Hain  
REGISTRATION NUMBER: 24864  
REFERENCE/DOCKET NUMBER: 92H832  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 665 2400  
TELEFAX: (908) 771 6159  
TELEX: 219484  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Region  
LOCATION: 369..419  
OTHER INFORMATION: /note= "Alternative C-termini of  
OTHER INFORMATION: HSA(1-n)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..585  
OTHER INFORMATION: /note= "Amino acid sequence of  
OTHER INFORMATION: natural HSA"  
US-08-153-799-14  
Query Match 82.0%; Score 3103; DB 1; Length 585;  
Best Local Similarity 100.0%; Pred. No. 3.9e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDDNPNLPLVRPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOPEPERNECFLOHKDDNPNLPLVRPEV 120  
Qy 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFVAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180  
Qy 181 KLDELDEGRKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFKAEFAVSKLVTDLTK 240  
Db 181 KLDELDEGRKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFKAEFAVSKLVTDLTK 240  
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVENDEMPA 300  
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECCCKPILKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYISVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYISVLLRLAKTYETTLK 360  
Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQGEYKFNALLVRYTKVPQVST 420  
Db 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCSELFQGEYKFNALLVRYTKVPQVST 420  
Qy 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLEHKEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVSRNLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLEHKEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALVEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVKHKPKAT 540  
Db 481 LVNRRPCFSALVEVDYVYKPEFNAETFTFHADICTLSEKEROIKQTALVELVKHKPKAT 540

```
QY 541 KEOLKAVMDDDFAAFVEKCKKADDDKTCFAEBEGKLVAAASQAALGL 585
Db 541 KEOLKAVMDDDFAAFVEKCKKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 4
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; CITY: 1020 First Avenue
; STATE: King of Prussia
; COUNTRY: Pennsylvania
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-WAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 82.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 3.9e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGFENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHKSVAHRFKDGLGFENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVNCTAFHDNEETFLKKLYEIIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEIIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180

QY 181 KLDELDEGKASAKQKLCASQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASAKQKLCASQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCKEPLLEKSHCHIAEVENDEMPA 300
Db 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCKEPLLEKSHCHIAEVENDEMPA 300

241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKECCKEPLLEKSHCHIAEVENDEMPA 300
QY 301 DLPSLAADFEVSKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLRLRAKTVETTTLEKC 360
Db 301 DLPSLAADFEVSKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLRLRAKTVETTTLEKC 360

QY 361 CAAADPHECYAKVDFDFKPLVBEPQNLIKQNCLELFEQLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAAADPHECYAKVDFDFKPLVBEPQNLIKQNCLELFEQLGEYKFNQALLVRYTKKVPQVST 420

QY 421 PTLVEVSRNLGVGSKCKHPEAKRMPCAEDYLSVVNLNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGVGSKCKHPEAKRMPCAEDYLSVVNLNOLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRFCFSALEVDETYVPEKFNAEFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540
Db 481 LVNRRFCFSALEVDETYVPEKFNAEFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540

QY 541 KEOLKAVMDDDFAAFVEKCKKADDDKTCFAEBEGKLVAAASQAALGL 585
Db 541 KEOLKAVMDDDFAAFVEKCKKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match 82.0%; Score 3103; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 3.9e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGFENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHKSVAHRFKDGLGFENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
```

```
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVS KLVDLTJK 240
Db 181 KLDELDRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVS KLVDLTJK 240
Qy 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPILLESKSHCTAEVNDMPA 300
Db 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPILLESKSHCTAEVNDMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360
Qy 361 CAAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLCVLHEKTPVSDRVTCKCTES 480
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 540
Qy 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 6
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 086957CD1
US-09-976-594-977

Query Match 82.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
```

```
Db 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVS KLVDLTJK 240
Db 205 KLDELDRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVS KLVDLTJK 264
Qy 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPILLESKSHCTAEVNDMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECCCKPILLESKSHCTAEVNDMPA 324
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTTLEKC 384
Qy 361 CAAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 385 CAAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLCVLHEKTPVSDRVTCKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLCVLHEKTPVSDRVTCKCTES 504
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVKHKPKAT 564
Qy 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 565 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 7
US-09-919-039-370
; Sequence 370, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370

Query Match 82.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-84;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVS KLVDLTJK 240
Db 205 KLDELDRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEFAEVS KLVDLTJK 264
```

QY	241	VHTECHGDLLECA	DRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA	300
DB	265	VHTECHGDLLECA	DRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA	324
QY	301	DLPSLAADFVESK	OVCKNYAKAOKVFLGMFLYEVYARSHPDYSVWLLRLAKTVEYTTLEK	360
DB	325	DLPSLAADFVESK	OVCKNYAKAOKVFLGMFLYEVYARSHPDYSVWLLRLAKTVEYTTLEK	384
QY	361	CAAADPHECYAK	VFDEFKPLVEEPPQNLKQNCSELFQOLGEYKQFQNALLVRYTKKVPQVST	420
DB	385	CAAADPHECYAK	VFDEFKPLVEEPPQNLKQNCSELFQOLGEYKQFQNALLVRYTKKVPQVST	444
QY	421	PTLVEVSRNLGH	KGSKCKHPKAPKMPCAEDYLSVWLNQLCVLHEKTPVSDRVTCKCTES	480
DB	445	PTLVEVSRNLGH	KGSKCKHPKAPKMPCAEDYLSVWLNQLCVLHEKTPVSDRVTCKCTES	504
QY	481	LVNRPKPSALEV	DEVTYYPKPFNAETFTFHADICTLSEKEROIKQOTALVELVHKHPKAT	540
DB	505	LVNRPKPSALEV	DEVTYYPKPFNAETFTFHADICTLSEKEROIKQOTALVELVHKHPKAT	564
QY	541	KSQLKAVMDDFA	AFVEKCKCKADDKETCPAEEGKKLVAASQAALGL	585
DB	565	KSQLKAVMDDFA	AFVEKCKCKADDKETCPAEEGKKLVAASQAALGL	609

RESULT 8  
 US-08-797-689-2  
 ; Sequence 2, Application US/08797689  
 ; Patent No. 5876969  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fleer, Reinhard  
 ; APPLICANT: Fournier, Alain  
 ; APPLICANT: Guillon, Jean-Dominique  
 ; APPLICANT: Jung, Gérard  
 ; APPLICANT: Yeh, Patrice  
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
 ; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
 ; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 ; STREET: 500 Arcola Road, 3C43  
 ; CITY: Collegeville  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19426  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: Macintosh  
 ; OPERATING SYSTEM: System 7.1  
 ; SOFTWARE: Word 5.1 (PatentIn)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/797,689  
 ; FILING DATE: 31-JAN-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/256,927  
 ; FILING DATE: 28-JUL-1994  
 ; APPLICATION NUMBER: FR 92/01064  
 ; FILING DATE: 31-JAN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/FR93/00085  
 ; FILING DATE: 28-JAN-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Smith Ph.D., Julie K.  
 ; REGISTRATION NUMBER: P-38,619  
 ; REFERENCE/DOCKET NUMBER: ST92006-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (610) 454-3839  
 ; TELEFAX: (610) 454-3808  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:



MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith P.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2

Query Match 82.0%; Score 3103; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 4.2e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDGLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 25 DAHSEVAHFKDGLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

Qy 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAFVSVKLVTDLTK 240  
Db 205 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAFVSVKLVTDLTK 264

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 384

Qy 361 CAADPHCEYAKVDFEPEKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPOVST 420  
Db 385 CAADPHCEYAKVDFEPEKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPOVST 444

Qy 421 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHKEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHKEKTPVSDRVTKCCTES 504

Qy 481 LVNRRPCFSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 564

Qy 541 KEQLKAVMDDFAAAFVEKCKCADDKCTCFABEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDDFAAAFVEKCKCADDKCTCFABEGKKLVAASQAALGL 609

RESULT 10  
US-09-949-016-11170  
; Sequence 11170, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11170  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11170

Query Match 82.0%; Score 3103; DB 4; Length 622;  
Best Local Similarity 100.0%; Pred. No. 4.3e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKDGLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 38 DAHSEVAHFKDGLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 97

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 98 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 157

Qy 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 158 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 217

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAFVSVKLVTDLTK 240  
Db 218 KLDELDEGKASSAKORLKASLQKFGERAFKAWAVARLSORFPKAFVSVKLVTDLTK 277

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300  
Db 278 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 337

Qy 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360  
Db 338 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 397

Qy 361 CAADPHCEYAKVDFEPEKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPOVST 420  
Db 398 CAADPHCEYAKVDFEPEKPLVEEPONLIKQNCLEFQOLGEYKFNALLVRYTKVPOVST 457

Qy 421 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHKEKTPVSDRVTKCCTES 480  
Db 458 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHKEKTPVSDRVTKCCTES 517

Qy 481 LVNRRPCFSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540  
Db 518 LVNRRPCFSALEVDVETYPKEFNAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 577

Qy 541 KEQLKAVMDDFAAAFVEKCKCADDKCTCFABEGKKLVAASQAALGL 585  
Db 578 KEQLKAVMDDFAAAFVEKCKCADDKCTCFABEGKKLVAASQAALGL 622

RESULT 11  
US-08-256-938-4  
; Sequence 4, Application US/08256938  
; Patent No. 5665863  
; GENERAL INFORMATION:  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE  
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.0 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,938  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92/01065  
; FILING DATE: 31-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goodman, Rosanne  
; REGISTRATION NUMBER: 32,534  
; REFERENCE/DOCKET NUMBER: ST92007-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3817  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-256-938-4  
Query Match 82.0%; Score 3103; DB 1; Length 787;  
Best Local Similarity 100.0%; Pred. No. 6.5e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHFRKDLGFENFALVLIAPAOYLQCPPEHVKLVNEVTEFAKTCVADESAE 60  
DB 203 DAHSEVAHFRKDLGFENFALVLIAPAOYLQCPPEHVKLVNEVTEFAKTCVADESAE 262  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 322  
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYVAPELLFPAKYKAAFTCCQAAADKAAACLLP 180  
DB 323 DVNCTAFHNEETFLKYLVEIARRHPYFYVAPELLFPAKYKAAFTCCQAAADKAAACLLP 382  
QY 181 KLDELDEGKASAKORLKCASQKQGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240  
DB 383 KLDELDEGKASAKORLKCASQKQGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 442  
QY 241 VHTTECHGDLLECDRADRLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 443 VHTTECHGDLLECDRADRLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 502  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMPFLYIYARRHPDYSVVLRLAKTYETTLK 360  
DB 503 DLPSLAADFVESKDVCKNVAEAKDVLGMPFLYIYARRHPDYSVVLRLAKTYETTLK 562

QY 361 CAAADPHECYAKVFDEFKELVEEPQNLIKONCELFQQLGEYKPFQNALLVRYTKVPQVST 420  
DB 563 CAAADPHECYAKVFDEFKELVEEPQNLIKONCELFQQLGEYKPFQNALLVRYTKVPQVST 622  
QY 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
DB 623 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 682  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPKAT 540  
DB 683 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPKAT 742  
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585  
DB 743 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 787  
RESULT 12  
US-08-797-689-16  
; Sequence 16, Application US/08797689  
; Patent No. 5876969  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guitton, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D. Julie K.  
; REGISTRATION NUMBER: P-36,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-797-689-16

Query Match 82.0%; Score 3103; DB 2; Length 787;  
Best Local Similarity 100.0%; Pred. No. 6.5e-84;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKOLGEENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db  
Qy 203 DAHSEVAHFKOLGEENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 262  
Db  
Qy 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 263 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 322  
Qy 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCCAADKAACLLP 180  
Db 323 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCCAADKAACLLP 382  
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKABFAEVSCLVTDLT 240  
Db 383 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKABFAEVSCLVTDLT 442  
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300  
Db 443 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 502  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVVLLRLAKTYETTTLEKC 360  
Db 503 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVVLLRLAKTYETTTLEKC 562  
Qy 361 CAAADPHECVAKVDFEFPKPLVEEPQNLIKQNCFLGEYKFNQALLVRYTKKVPQVST 420  
Db 563 CAAADPHECVAKVDFEFPKPLVEEPQNLIKQNCFLGEYKFNQALLVRYTKKVPQVST 622  
Qy 421 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 623 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 682  
Qy 481 LVNRRPCFSALVDETVYPKFNFAETFTFHADICTLSEKERQIKKQALVELVXHKPKAT 540  
Db 683 LVNRRPCFSALVDETVYPKFNFAETFTFHADICTLSEKERQIKKQALVELVXHKPKAT 742  
Qy 541 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEBGKKLVAASQAALGL 585  
Db 743 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEBGKKLVAASQAALGL 787

RESULT 13  
US-09-984-186-16  
; Sequence 16, Application US/09984186  
; Patent No. 6866179  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guittou, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/984,186  
; FILING DATE: 29-Oct-2001  
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-984-186-16

Query Match 82.0%; Score 3103; DB 4; Length 787;  
Best Local Similarity 100.0%; Pred. No. 6.5e-84;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFKOLGEENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60  
Db 203 DAHSEVAHFKOLGEENFKALVLIAPAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 262  
Qy 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 263 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 322  
Qy 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCCAADKAACLLP 180  
Db 323 DVMTAFHDNEETFLKKYLYEIARRHPYFYVAPELLFFAKRYKAAFTCCCAADKAACLLP 382  
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKABFAEVSCLVTDLT 240  
Db 383 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKABFAEVSCLVTDLT 442  
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300  
Db 443 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 502  
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVVLLRLAKTYETTTLEKC 360  
Db 503 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVVLLRLAKTYETTTLEKC 562  
Qy 361 CAAADPHECVAKVDFEFPKPLVEEPQNLIKQNCFLGEYKFNQALLVRYTKKVPQVST 420  
Db 563 CAAADPHECVAKVDFEFPKPLVEEPQNLIKQNCFLGEYKFNQALLVRYTKKVPQVST 622  
Qy 421 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 623 PTLVEVSRLNGKVGSKCKGHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 682  
Qy 481 LVNRRPCFSALVDETVYPKFNFAETFTFHADICTLSEKERQIKKQALVELVXHKPKAT 540  
Db 683 LVNRRPCFSALVDETVYPKFNFAETFTFHADICTLSEKERQIKKQALVELVXHKPKAT 742  
Qy 541 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEBGKKLVAASQAALGL 585  
Db 743 KEQLKAVMDPFAAFVEKCKCCKADDETCFAEBGKKLVAASQAALGL 787

RESULT 14  
US-08-222-619-3  
; Sequence 3, Application US/08222619

Patent No. 5652352  
GENERAL INFORMATION:  
APPLICANT: Lichenstein, Henri  
APPLICANT: Lyons, David  
APPLICANT: Wurfel, Mark  
APPLICANT: Wright, Samuel  
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
TITLE OF INVENTION: Protein  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Center, Patent Operations/RR  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: U.S.  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,619  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-222-619-3

Query Match 81.9%; Score 3099; DB 1; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 180  
DB 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 204  
QY 181 KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 264  
QY 241 VHTCCGHDLLCADDRAIDLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 300  
DB 265 VHTCCGHDLLCADDRAIDLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYIYARRHPDYSVVLRLAKTYETTLK 360  
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYIYARRHPDYSVVLRLAKTYETTLK 384  
QY 361 CAAADPHCEYAKVDFEKPPIVEEPONLIKONCELPQOLGEYKQNALVRYTKVPQVST 420  
DB 385 CAAADPHCEYAKVDFEKPPIVEEPONLIKONCELPQOLGEYKQNALVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDVETVVPKFEFNAETFFHADICTISEKEROIKKQOTALVELVKHPKAT 540  
DB 505 LVNRRPCFSALEVDVETVVPKFEFNAETFFHADICTISEKEROIKKQOTALVELVKHPKAT 564

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585  
DB 565 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609  
RESULT 15  
US-08-433-037-4  
Sequence 4, Application US/08433037  
Patent No. 5707828  
GENERAL INFORMATION:  
APPLICANT: Sreekrishna, Kotikanyadan  
APPLICANT: Barr, Kathryn A.  
APPLICANT: Brierley, Russell A.  
APPLICANT: Thill, Gregory P.  
APPLICANT: Tschopp, Jueg P.  
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
TITLE OF INVENTION: PICHIA PASTORIS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9108Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 81.9%; Score 3099; DB 1; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHRFKDLGEBNFKALVLIAPAOYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 180  
DB 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 204  
QY 181 KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240  
DB 205 KLDELDEGKASSAKORLKCASLOKTFGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 264  
QY 241 VHTCCGHDLLCADDRAIDLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 300

Db 265 VHTCCGHDLLCADDADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVNDMPA 324  
Qy 301 DLPSLAADFVSKVKYAEAKDVFGLGMFLYAYARRHPDYSVLLLRLLAKTYETTLK 360  
Db 325 DLPSLAADFVSKVKYAEAKDVFGLGMFLYAYARRHPDYSVLLLRLLAKTYETTLK 384  
Qy 361 CAAADPHCYAKVDFEFPKPLVEEPONLIKQNCFLFEQLGEYKFNALLVRYTKVPOVST 420  
Db 385 CAAADPHCYAKVDFEFPKPLVEEPONLIKQNCFLFEQLGEYKFNALLVRYTKVPOVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKMPCADYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKMPCADYLSVNLQCLVLEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 564  
Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 585  
Db 565 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 609

## RESULT 16

US-08-897-956A-2  
; Sequence 2, Application US/08897956A  
; Patent No. 6423512  
; GENERAL INFORMATION:  
; APPLICANT: Mary Ellen Digan  
; APPLICANT: Philip Lake  
; APPLICANT: Hermann Gram  
; TITLE OF INVENTION: Fusion Polypeptides  
; FILE REFERENCE: 600-7244/CPA  
; CURRENT APPLICATION NUMBER: US/08/897,956A  
; CURRENT FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/022,689  
; PRIOR FILING DATE: 1996-07-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 81.9%; Score 3099; DB 4; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYQLQQCFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHRPKDLGEENFKALVLIAPAYQLQQCFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDNDPMLPRVLRPEV 120  
Db 85 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDNDPMLPRVLRPEV 144  
Qy 121 DVMTCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 145 DVMTCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204  
Qy 181 KLDELDRGKASSAKQRLKCSLQKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDRGKASSAKQRLKCSLQKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLT 264  
Qy 241 VHTCCGHDLLCADDADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVNDMPA 300  
Db 265 VHTCCGHDLLCADDADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVNDMPA 324  
Qy 301 DLPSLAADFVSKVKYAEAKDVFGLGMFLYAYARRHPDYSVLLLRLLAKTYETTLK 360  
Db 325 DLPSLAADFVSKVKYAEAKDVFGLGMFLYAYARRHPDYSVLLLRLLAKTYETTLK 384

Qy 361 CAAADPHCYAKVDFEFPKPLVEEPONLIKQNCFLFEQLGEYKFNALLVRYTKVPOVST 420  
Db 385 CAAADPHCYAKVDFEFPKPLVEEPONLIKQNCFLFEQLGEYKFNALLVRYTKVPOVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKMPCADYLSVNLQCLVLEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKMPCADYLSVNLQCLVLEKTPVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540  
Db 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 564  
Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 585  
Db 565 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEKGKLVAAASQAALGL 609

## RESULT 17

PCT-US95-04075-3  
; Sequence 3, Application PC/TUS9504075  
; GENERAL INFORMATION:  
; APPLICANT: AMGEN INC.  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RRC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04075  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US95-04075-3

Query Match 81.9%; Score 3099; DB 5; Length 609;  
Best Local Similarity 99.8%; Pred. No. 5.5e-84;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYQLQQCFEDHVKLVNEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHRPKDLGEENFKALVLIAPAYQLQQCFEDHVKLVNEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDNDPMLPRVLRPEV 120  
Db 85 NCDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDNDPMLPRVLRPEV 144  
Qy 121 DVMTCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180  
Db 145 DVMTCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 204  
Qy 181 KLDELDRGKASSAKQRLKCSLQKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDRGKASSAKQRLKCSLQKFGERAFAKAWARLSORFPKAEFAEVSCLVTDLT 264  
Qy 241 VHTCCGHDLLCADDADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVNDMPA 300  
Db 265 VHTCCGHDLLCADDADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVNDMPA 324

QY 301 DLPSLAADFVESKDVCKNVAEADVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVESKDVCKNVAEADVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 384  
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGKYKFNQALLVRYTKKVPQVST 420  
DB 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGKYKFNQALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHPKAT 540  
DB 505 LVNRRPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHPKAT 564  
QY 541 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB 565 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 609

RESULT 18

US-08-448-196A-3  
; Sequence 3, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/WFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-08-448-196A-3

Query Match 81.7%; Score 3093; DB 1; Length 585;  
Best Local Similarity 99.7%; Pred. No. 7.7e-84;  
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGKGFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
DB 1 DAHSEVAHRFKDLGKGFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEYARRHPYFYAPBLLFFAKRYKAAFTTECCQAADKAAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEYARRHPYFYAPBLLFFAKRYKAAFTTECCQAADKAAACLLP 180  
QY 181 KLDELDEGKASAKORLKCASLQKFGERAFAKAWAVARLSQRPPKAEFAEVSKLVDLTG 240  
DB 181 KLDELDEGKASAKORLKCASLQKFGERAFAKAWAVARLSQRPPKAEFAEVSKLVDLTG 240  
QY 241 VHTTECCGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEYENDEMPA 300  
DB 241 VHTTECCGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEYENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNVAEADVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 360  
DB 301 DLPSLAADFVESKDVCKNVAEADVFLGMFLYFYARRHPDYSVVLRLRLAKTYETTTLEKC 360  
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGKYKFNQALLVRYTKKVPQVST 420  
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKQNCLEFQOLGKYKFNQALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHPKAT 540  
DB 481 LVNRRPCFSALEVDYETVVPKEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHPKAT 540  
QY 541 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585  
DB 541 KEOLKAVMDDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 19

US-08-984-176-1  
; Sequence 1, Application US/08984176  
; Patent No. 5948609  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C  
; APPLICANT: HO, JOSEPH X  
; APPLICANT: RUKER, FLORIAN  
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
; FILE REFERENCE: 08/984,176  
; CURRENT APPLICATION NUMBER: US/08/984,176  
; CURRENT FILING DATE: 1997-12-03  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-984-176-1  
Query Match 81.7%; Score 3093; DB 2; Length 585;  
Best Local Similarity 99.7%; Pred. No. 7.7e-84;  
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDLGKGFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
DB 1 DAHSEVAHRFKDLGKGFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120  
QY 121 DVNCTAFHDNEETFLKKYLYEYARRHPYFYAPBLLFFAKRYKAAFTTECCQAADKAAACLLP 180  
DB 121 DVNCTAFHDNEETFLKKYLYEYARRHPYFYAPBLLFFAKRYKAAFTTECCQAADKAAACLLP 180

Db 121 DVMCTAFHNEETFLKKYVEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Qy 181 KLDELDEGRASAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEVSRLVTLDTK 240  
Db 181 KLDELDEGRASAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEVSRLVTLDTK 240  
Qy 241 VHTCCGHDLLLECCADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGHDLLLECCADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYARAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYARAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPPQNLIKQNCSELFQKQFONALLVRYTKVPQVST 420  
Db 361 CAAADPHCEYAKVDFEFPKPLVEEPPQNLIKQNCSELFQKQFONALLVRYTKVPQVST 420  
Qy 421 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQALVELVHKPKAT 540  
Db 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQALVELVHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDETCFAEKGKLVAAASQAALGL 585  
Db 541 KEQLKAVMDDFAAFVEKCKCCKADDETCFAEKGKLVAAASQAALGL 585  
RESULT 20  
US-08-448-196A-5  
; Sequence 5, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal

US-08-448-196A-5  
Query Match 65.0%; Score 2450.9; DB 1; Length 583;  
Best Local Similarity 75.8%; Pred. No. 4.6e-65;  
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;  
Qy 1 DAKSEVAHFKDLGRNFKALVLIAPAOYLQCCPPREDHVKLVNEVTEFAKTCVADSEAE 60  
Db 1 DTHKSEIAHFNLDGEXHFKGLVLAFAQYLOQCPFPEDHVKLVNEVTEFAKTCVADSEAE 60  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
Qy 121 DVMCTAFHNEETFLKKYVEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Db 121 DVMCTAFHNEETFLKKYVEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAAACLLP 180  
Qy 181 KLDELDEGRASAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEVSRLVTLDTK 240  
Db 181 KLDELDEGRASAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEVSRLVTLDTK 240  
Qy 241 VHTCCGHDLLLECCADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Db 241 VHTCCGHDLLLECCADDRADLAKYICENQDSISSKLECCCKPDLLEKSHCIAEVENDEMPA 300  
Qy 301 DLPSLAADFVESKDVCKNYARAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Db 301 DLPSLAADFVESKDVCKNYARAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360  
Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPPQNLIKQNCSELFQKQFONALLVRYTKVPQVST 420  
Db 361 CAAADPHCEYAKVDFEFPKPLVEEPPQNLIKQNCSELFQKQFONALLVRYTKVPQVST 420  
Qy 421 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 421 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Qy 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQALVELVHKPKAT 540  
Db 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQALVELVHKPKAT 540  
Qy 541 KEQLKAVMDDFAAFVEKCKCCKADDETCFAEKGKLVAAASQAALGL 583  
Db 541 KEQLKAVMDDFAAFVEKCKCCKADDETCFAEKGKLVAAASQAALGL 583  
RESULT 21  
US-08-448-196A-4  
; Sequence 4, Application US/08448196A  
; Patent No. 5780594  
; GENERAL INFORMATION:  
; APPLICANT: CARTER, DANIEL C.  
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
; TITLE OF INVENTION: RELATED PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NASA  
; STREET: MARSHALL SPACE FLIGHT CENTER  
; CITY: HUNTSVILLE  
; STATE: ALABAMA  
; COUNTRY: USA  
; ZIP: 35812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,196A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROAD JR., ROBERT L.  
; REGISTRATION NUMBER: 18,757  
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 205-544-0021  
; TELEFAX: 205-544-0258  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal

```
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-4

Query Match 64.8%; Score 2450.9; DB 1; Length 583;
Best Local Similarity 75.8%; Pred. No. 7.9e-65;
Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFDLGBENFKALVLIAPAOYLQCCPDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DTHKSEIAHRFDLGBEHFKGLVLIASFQYLQCCPDHVKLVNELTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
DB 61 GCEKSLHTLFGDELCKVASLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFPKRYKAAFTCCQAADKACLLP 180
DB 120 NTLCDFEKADEKKFWGKLYEIVARRHPYFYAPPELLFYANKYNGVFQCCQAEKACLLP 179
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPKAEFAVSKLVTDLTK 240
DB 180 KIETREKVLASSARQELRCASLOKFGERALKAWAVARLSORFPKAEFAVSKLVTDLTK 239
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDISSKLKECCBKPLLEKSHCIAEVENDEMPA 300
DB 240 VHKECCHGDLLECCADRADLAKYICENQDISSKLKECCBKPLLEKSHCIAEVEKDAIPE 299
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYARRHPDYSVLLRLAKTYETTLK 360
DB 300 NLPLTADFAEDKDVCKNYQEAADFLGSLFYYSRRHPYAVSVLLRLAKEYEATLEEC 359
QY 361 CAADAPHECYAKVDFEKPVLVEEPQNLIKQNCBELFQOLGEYKFGQNALIVRYTRKVPQVST 420
DB 360 CAKDDPHACYATVDFDKLHLVDEPQNLIKQNCDOFEKLGEGYQNALIVRYTRKVPQVST 419
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 420 PTLVEVSRNLGKVGTRCCTKPESERMPCETDYLSLINRLCVLHEKTPVSEKVTCKCTES 479
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKEROIKQOTALVELVGHKPKAT 540
DB 480 LVNRRPCFSALTPTDETVPVKAPEKLTFTFHADICTLPTDEKQIKQOTALVELLKHKPKAT 539

541 KEOLKAWMDFAAFVEKCCKADKDETCFAEGEKKLVAASQAAL 583
540 EEQLKTWENFVAFVDKCCAADKKEACFAVEGPKLVVSTOTAL 582

RESULT 22
US-08-448-196A-6
; Sequence 6, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-6

Query Match 64.3%; Score 2432.9; DB 1; Length 583;
Best Local Similarity 75.0%; Pred. No. 2.7e-64;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFDLGBENFKALVLIAPAOYLQCCPDHVKLVNEVTEFAKTCVADES 60
DB 1 DTHKSEIAHRFDLGBENFKGLVLIASFQYLQCCPDHVKLVNELTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 120
DB 61 GCDKSLHTLFGDELCKVASLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVRPEV 119
QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFPKRYKAAFTCCQAADKACLLP 180
DB 120 DTLCAEFKADEKKFWGKLYEIVARRHPYFYAPPELLFYANKYNGVFQCCQAEKACLLP 179
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORFPKAEFAVSKLVTDLTK 240
DB 180 KIDAMREKVLASSARQELRCASLOKFGERALKAWAVARLSORFPKAEFAVSKLVTDLTK 239
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDISSKLKECCBKPLLEKSHCIAEVENDEMPA 300
DB 240 VHKECCHGDLLECCADRADLAKYICENQDISSKLKECCBKPLLEKSHCIAEVEKDAIPE 299
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLFYARRHPDYSVLLRLAKTYETTLK 360
DB 300 NLPLTADFAEDKDVCKNYQEAADFLGSLFYYSRRHPYAVSVLLRLAKEYEATLEEC 359
QY 361 CAADAPHECYAKVDFEKPVLVEEPQNLIKQNCBELFQOLGEYKFGQNALIVRYTRKVPQVST 420
DB 360 CAKDDPHACYATVDFDKLHLVDEPQNLIKQNCDOFEKLGEGYQNALIVRYTRKVPQVST 419
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 420 PTLVEVSRNLGKVGTRCCTKPESERMPCETDYLSLINRLCVLHEKTPVSEKVTCKCTES 479
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKEROIKQOTALVELVGHKPKAT 540
DB 480 LVNRRPCFSDLTDETVPVKAPEKLTFTFHADICTLPTDEKQIKQOTALVELLKHKPKAT 539
```







PCT-US95-04075-4

Qy	4	KSEVAHR	-----FKOLGSENFKALVLIAPQYLQOCFFDFHVKLVNEVTEFAKT	52
Db	17	ESRILHRNEYGIASILDVSVOCTAEISLADLATIFFAQFVQEATYKEVSSQWDXDALTAIEK	76	
Qy	53	CVADESABENDKSLHTLFGDKLCTVATUARETVGEMADCCAKOPEBNEBCFLOHKDDNP-N	111	
Db	77	PTGDEQSSGCLLENQOLPAFLEELCHEKEILEKYGH--SDCSQSEGRHNCFLAHKPTPAS	135	
Qy	112	LPRLVRPVDVMCTAFPHDNEETFLKKYLYEIRARRHPYFYVAPELLFPAKEYKAAFTCCCOA	171	
Db	136	IPLFQVPEPVTSCSEAYBEDRETFMNKFYIEIRARRHPFLYAPTILLWAARYDKIILPSCCKA	195	
Qy	172	ADKAACLLPKLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPPKAEFAEV	231	
Db	196	ENAVECFQTKAATVTKELRESSLLNQHACAVMNFGRTRFQAITVTKLSQKTKYNFTEI	255	
Qy	232	SKLVTDLTKVATECHGDLLDECADRDADIACYICENQDSISSKLEKCECKPILLEKSHCIA	291	
Db	256	QKVLVDVAHVHEHCCRGDVLDCLODGEKIMSYICSQDITLSNKITCECKLTTLERGQCII	315	
Qy	292	EVENDEMPADLPSLAADFVESKDVKCNVYABAKDVLGMLFELYARRHPDYSVVLILRLAK	351	
Db	316	HAENDEKEGLSPNLRNFIJGDRDFNQFSSGSEKNIFLASFVHEYSRRHPQLAVSVILRVAK	375	
Qy	352	TYETTLKCCAAADPHECVAKYFDEPKPLVEBPQNLIKONCELFEOLGEYKFNALLVRY	411	
Db	376	GYOELLKCFQTFENPLECODKGEELQKYIQBSQALAKRSCGLFQKLGELYONAFVAY	435	
Qy	412	TKKVPQVSTPTLVEYSRNLKGKVGSKCKHPEAKRMPECAEDYLSVNLNQLCVLHEKTPVSD	471	
Db	436	TKKAPQLTSSELWATRMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEMTPVNP	495	
Qy	472	RVTKCCTBSLNRNRPCCFSALEVDETVYVPKEFNAETTFTHADICTLSEKEROIKKQATLVE	531	
Db	496	GVGQCTCTSYANRRPCFSSLVDETVVPAFSDDKIFHFKDLQCAQGVALQTMKQSFLLN	555	
Qy	532	LVKHKPKATKEQIKAVMDDFAPAVEKCKKADDKETCFABEGKLVAAASONALGL	585	
Db	556	LVKOKFOITEEOLEAVIADFSGLEKCCQGOEVCVFABEGKLI1KTRAAALGV	609	

532 LVKHKPKATKEQLKAVMDDFAAFVEKCKKADDTCTFAEEGKLVAAASAALGL 585  
QY  
556 LVKCKPOTTEGQLEAVTADNESHGILEKCCGCGEOEVCFAEEGKLVKSTRAALGV 609  
DQ

Query Match 33.1%; Score 1252.7; DB 4; Length 612;

```
Best Local Similarity 39.7%; Pred. No. 2.6e-29;
Matches 236; Conservative 114; Mismatches 231; Indels 13; Gaps 3;

QY 4 KSEVAHR-----FKDLGEENFKALVLIARFAQYLQCCPFEDHVKLVNEVTEFAKT 52
Db
20 ESRTLHRNEYGIASILDYSQCTAEISLADLATIFFAQVQOEATYKEVSKXVKDALTAEIK 79
QY 53 CVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNP-N 111
Db
80 PTGDEOSSGLENQLPAFLBELCHEKEILEIKYGH-SDCCSQSEGRHNCFLAHHKKTTPAS 138
QY 112 LPLRVPEVDVMTAFHDNEETFLKKYLYEIARRHPFYFAPPELLFFAKRYKAAFTCCQA 171
Db
139 IPLFQVPEPVTSCAEYEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKIIPSCCKA 198
QY 172 ADKAAACLLPKDLRLDEGRKASSAKQRLKCSLQKFGERAFKAWARLSQRPFKAEFAV 231
Db
199 ENAVECFQTKAAATVTKELRESSLLNQHACAVMKNGFTRTFQAITVTKLSQKFTKVNFTBI 258
QY 232 SKLVTDLTKVHTCCCHGDLLEACADDRADLAKYICENQDSTSSKLBKCCCKPLLEKSHCIA 291
Db
259 QKLVLDVAHVHEHCGRGVDLCLQDGEKIMSYICSQODTLSNKITECCCKLTTLERGCQII 318
QY 292 EVNDEMPADLPSSAADFVESKDVCKNYABAKDVLGMLFYEYARRHPDYSVVLLRLAK 351
Db
319 HAENDEKPEGLSPNLNRFGLDRDFNQFSGGEKNIFLASFVHEYSSRRHPQLAVSVILRVAK 378
QY 352 TYETTLKCCAAADPHCEYAKVDFEKPVLVEEPQNLKONCELFEQOLGEYKFFQNALIVRY 411
Db
379 GYOELLEKCFQTNPLECQDKGEEELQKYIQESQALAKRSCGLFQKLGYYLQNEFLVAY 438
QY 412 TKKVPQVSTPTLVEVSRNLGKVGSKCKCHPEAKRMPCAEVDYLSVLNLQCLVLHEKTPVSD 471
Db
439 TKKAPOLTSSELMAITRKAAATAATCCQLSEDKLLACGEGAADIIGHLICIRHEMTVPNP 498
QY 472 RVTKKCTESLVNRRPCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVE 531
Db
499 GVGQCCTSSYANRRPCFSSLVVDVETVYPPAFSDDKFIHFKDLCAQOQVALQTMKQBFLIN 558
QY 532 LVKHGKPKATKEQLKAVMDDFAAFEVKCKKADDKETCFABEGKKLVAASQAALGL 585
Db
559 LVKQKQPQITEOLEALVIADFSGLLEKCCQGEQEVCFABEGQKLIISKTRAAALGV 612

RESULT 29
US-09-186-949A-2
; Sequence 2, Application US/09186949A
; Patent No. 6416734
; GENERAL INFORMATION:
; APPLICANT: Murgita, Robert A.
; TITLE OF INVENTION: Recombinant Alpha-Fetoprotein For
; TITLE OF INVENTION: Treating and Diagnosing Cancers
; FILE REFERENCE: 06727/004002
; CURRENT APPLICATION NUMBER: US/09/186,949A
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 08/758,757
; PRIOR FILING DATE: 1996-12-03
; PRIOR APPLICATION NUMBER: US 08/377,311
; PRIOR FILING DATE: 1995-01-24
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-186-949A-2

Query Match 32.0%; Score 1211.3; DB 4; Length 609;
Best Local Similarity 38.8%; Pred. No. 4.3e-28;
Matches 231; Conservative 113; Mismatches 235; Indels 17; Gaps 4;

QY 4 KSEVAHR-----FKDLGEENFKALVLIARFAQYLQCCPFEDHVKLVNEVTEFAKT 52
Db

Query Match 32.0%; Score 1210.9; DB 2; Length 590;
Best Local Similarity 38.9%; Pred. No. 4.2e-28;
Matches 230; Conservative 115; Mismatches 235; Indels 11; Gaps 4;

QY 3 HKSE-----VAHRFKDLGLENFKALVLIARFAOYLQCCPFEDHVKLVNEVTEFAKTCAVE 57
Db
3 HRNEYGIASILDYSQCTAEISLADLATIFFAQVQOEATYKEVSKXVKDALTAEIKETGDE 62
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNP---LPR 114
Db
63 QSSGLENQLPAFLBELCHEKEILEIKYGH-SDCCSQSEGRHNCFLAHHKKTTPAS 119
QY 115 LVRPEVDVMTAFHDNEETFLKKYLYEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADK 174
Db
120 FQVPEPVTSCAEYEDRETFMNKFIYEIARRHPFLYAPTILLWAARYDKIIPSCCKAENA 179
```

Qy	175	AACLLPKLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFPKAEFAEVS	234
Db	180	VECFQTKAAVTWKELRESSLLNQHACPMKNFGTRTFQAITVTKLSQKFTKVNFT	239
Qy	235	VTDLTKVHTCCCHGDLLECADDRADLAKYICENQDSISSKKECCEKPLLEKSHCIAE	294
Db	240	VLDVAHVHEHCCRADVLDCLQDGEKIMSYICSQQDTLSNKTCECKLTTLERGQCI	299
Qy	295	NDEMPADLPSLADPFVESKDVCKNYAEAKDVFLGMPLYEYARRHPDYSVVLLRL	354
Db	300	NDEKPEGLSPNLRNFLGDRDFNQSSGCKNIFLASFVHEYRRHPQLAVSVILEV	359
Qy	355	TLEKCCAAADPHCYAKVDFEPKPLVEEPQNLKQNCSELFQELGEYKFNALLV	414
Db	360	ELLEKCFQTNPLEFCQDGBEELQKYIQESQALAKRSCGLFQKLGEYYLQNEFL	419
Qy	415	VPQVSTFTLVEVSRNLKVGSKCKHPEAKMPCAEYLSVVLNQLCVLHEKTPV	474
Db	420	APQLTSSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIIGHLCIRHE	479
Qy	475	KCTESLVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKQ	534
Db	480	QCTSSYANRRPCFSSLVVDVETVVPAPFSDDKFIHFKDLCOAQGVALQRMK	539
Qy	535	HKPKATKEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL	585
Db	540	KPQITEQLEALIAFDSGLLEKCCQGEQEVCFABEGQKLISKTGAAALGV	590

Search completed: October 17, 2005, 08:29:42  
Job time : 47 secs

**THIS PAGE BLANK (08/17/10)**

Result No.	Score	Query Match	Length	DB ID	Description
1	3156.5	83.4	763	15	US-10-609-346-2
2	3153.4	83.3	838	17	US-10-775-180-590
3	3152.4	83.3	838	18	US-10-775-204-1591
4	3152.8	83.3	876	17	US-10-775-180-82
5	3152.8	83.3	876	18	US-10-775-204-236
6	3152.3	83.3	819	17	US-10-775-180-603
7	3152.3	83.3	819	17	US-10-775-180-605
8	3152.3	83.3	819	18	US-10-775-204-1614
9	3152.3	83.3	819	18	US-10-775-204-1616
10	3152.2	83.3	794	18	US-10-775-204-1657
11	3152.2	83.3	809	18	US-10-775-204-1548





```
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 240
Db 205 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLM----- 586
Db 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMDLLQFLAFLVLL 624
QY 587 -----SPRLEV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKRRBQLALKNLAQINDIHOQYKILDVMLKGLFK 684

RESULT 3
US-10-775-204-1591
; Sequence 1591, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; PRIOR FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR APPLICATION NUMBER: 60/360,000
```

```
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 240
Db 205 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLM----- 586
Db 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMDLLQFLAFLVLL 624
QY 587 -----SPRLEV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKRRBQLALKNLAQINDIHOQYKILDVMLKGLFK 684

Query Match 83.3%; Score 3153.4; DB 18; Length 838;
Best Local Similarity 75.0%; Pred. No. 2.2e-75;
Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 240
Db 205 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLM----- 586
Db 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMDLLQFLAFLVLL 624
QY 587 -----SPRLEV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKRRBQLALKNLAQINDIHOQYKILDVMLKGLFK 684

US-10-775-204-1591
; TYPE: PRT
; ORGANISM: Homo sapiens
;
Query Match 83.3%; Score 3153.4; DB 18; Length 838;
Best Local Similarity 75.0%; Pred. No. 2.2e-75;
Matches 624; Conservative 15; Mismatches 37; Indels 156; Gaps 12;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPPEHVDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 180
Db 145 DVNCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQOAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 240
Db 205 KLDELDEGKASSAKQRLKASQKQGERAFKAWAVARLSQRPKPAFAEVSKLVTDLYK 264
QY 241 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENODSISKLKECCEKPLEKSHCHIAEVNDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYAYARRHPDYSVVLRLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLM----- 586
Db 565 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMDLLQFLAFLVLL 624
QY 587 -----SPRLEV----- 592
Db 625 SGMGATGTLRTSLDPSLEIYKKMFVKRRBQLALKNLAQINDIHOQYKILDVMLKGLFK 684

Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2222
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1591
LENGTH: 838
TYPE: PRT
ORGANISM: Homo sapiens
US-10-775-204-1591
```

QY 593 -----PCSHALPOGLSPGQVIVIRGLVQLQPKHFTVSLRDOAAHAPVTLRASFADR 643  
DB 685 VLEDSRTVLTADVLDPGFP-----QDEK-----LKDAFSHV-VENTAFPGDV 727  
QY 644 TLQ-----W-ISRWGOKKLIISAPFLFYQRFVLLFOEGGLK--LAIN 685  
DB 728 VLRFPRIHVHYFDHNSWNLLIRWG-----IS-----FCNQTGVFNQGHSPILSLM 774  
QY 686 GQGLG-----ATSMNQOAL-----BOLREL 706  
DB 775 AQELGISEKSNFQNPFKIDRTFIPSTDPFQKALREERKRKEIR 826

RESULT 4  
US-10-775-180-82  
; Sequence 82, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 82  
; LENGTH: 876  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-82

Query Match 83.3%; Score 3152.8; DB 17; Length 876;  
Best Local Similarity 70.4%; Pred. No. 2.5e-75;  
Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPQYLOQCPPEHVKLVNVEFPAKTCVADESAE 60  
DB 25 DAHSEVAHRFDLGEENFKALVLIAPQYLOQCPPEHVKLVNVEFPAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRV 144  
QY 121 DVNCTAFDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
DB 145 DVNCTAFDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204  
QY 181 KLDEURDEGKASAKORLKASLQKGFGERAFKAWAVARLSQRPFAEVSCLVTDLT 240  
DB 205 KLDEURDEGKASAKORLKASLQKGFGERAFKAWAVARLSQRPFAEVSCLVTDLT 264

QY 241 VHTTECHGDLLECCADRADLAKYICENODSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADPVEKDYCKNVAEAKDVLGMLFYEARHPDYVSVLLRLAKTYETTLEK 360  
DB 325 DLPSLAADPVEKDYCKNVAEAKDVLGMLFYEARHPDYVSVLLRLAKTYETTLEK 384  
QY 361 CAAADPHCYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYFQNALVRYTKKVPQVST 420  
DB 385 CAAADPHCYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYFQNALVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHKTVPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVKHKPKAT 540  
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQATALVELVKHKPKAT 564  
QY 541 KEQLXAVMDDFAAVFEKCKKADDDKTCFAEBEGKLVAAASQAALGLSPRLEV----- 592  
DB 565 KEQLXAVMDDFAAVFEKCKKADDDKTCFAEBEGKLVAAASQAALGL-----LEVAETPTYPW 620  
QY 593 -----PCSHALPOGLSPGQVIVIRGLVQLQPKHFT----- 622  
DB 621 RDAETGERLVCAQCPGTFVQRPCTRDSTTCGP-----CPRHYTQFWNYLER 669  
QY 623 -----VSLRDQAAHA----- 632  
DB 670 CRYCNVLGGEREEARACHATHNRACRKTGFFAHAGFCLEHASCPPGAGVIAPGTPSQN 729  
QY 633 -----PVTLRAS----- 639  
DB 730 TOQCPGPTFTFSASSSSSEQCQPHRNCNTALGLALNVPGSSSHDTLCTSGTGPPLSTRVPG 789  
QY 640 -----FADRTLWISR-----WG-----QKGLISAPFLFYPO 666  
DB 790 AECEERAVIDFVAFODISIKRLQRLLOALEAPEGHGPTPRAGRAALQLKL-----R 840  
QY 667 RFEVLLLFQEGGLKALNGQGLGATSMNQQAELQELRELIRSG 709  
DB 841 RRLTELLGAQDGALLVRL-----LQALRVARMGP 869

RESULT 5  
US-10-775-204-236  
; Sequence 236, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11

1 PRIOR APPLICATION NUMBER: 60/420,246  
2 PRIOR FILING DATE: 2002-10-23  
3 PRIOR APPLICATION NUMBER: 60/423,623  
4 PRIOR FILING DATE: 2002-11-05  
5 PRIOR APPLICATION NUMBER: 60/351,360  
6 PRIOR FILING DATE: 2002-01-28  
7 Remaining Prior Application data removed - See File Wrapper or PALM.  
8 NUMBER OF SEQ ID NOS: 2222  
9 SOFTWARE: Patentin Ver. 2.0  
10 SEQ ID NO 236  
11 LENGTH: 876  
12 TYPE: PRT  
13 ORGANISM: Homo sapiens  
14 US-10-775-204-236

Query Match 83.3%; Score 3152.8; DB 18; Length 876;  
Best Local Similarity 70.4%; Pred. No. 2.5e-75;  
Matches 622; Conservative 13; Mismatches 36; Indels 212; Gaps 10;

QY 1 DAHSEVAHRFKDGBENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDGBENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLPGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLPGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCOADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCOADKAACLLP 204  
QY 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLTK 240  
DB 205 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLTK 264  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVENDMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVENDMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYVARHPDYSVVLRLRLAKTVEITTELEK 360  
DB 325 DLPSLAADFVESKDVCKNVAEKDVLGMFLYVARHPDYSVVLRLRLAKTVEITTELEK 384  
QY 361 CAADPHCEYAKVDFEKLPIVEEPQNLIKONCELFQGLYKQFONALLVRYTKVPQVST 420  
DB 385 CAADPHCEYAKVDFEKLPIVEEPQNLIKONCELFQGLYKQFONALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKSGCKKHPEAKRMPCABDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKSGCKKHPEAKRMPCABDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVTVPKFENAEFTTFHADICTLSEKEROIKKOTALVELVKGHPKAT 540  
DB 505 LVNRRPCFSALEVDVTVPKFENAEFTTFHADICTLSEKEROIKKOTALVELVKGHPKAT 564  
QY 541 KEQLKAWMDPFAAFVEKCCADKDETCAFEAGKGLVAASQAALGMSPLREV----- 592  
DB 565 KEQLKAWMDPFAAFVEKCCADKDETCAFEAGKGLVAASQAALG-----LEVAETPTYPW 620  
QY 593 -----PCSHALPQGLSPGQVILVRLGLVQBPQKHT----- 622  
DB 621 RDAETGERLVCAQCPGFTVQRCRDSPTTCGE-----CPRHYTQFNWYLER 669  
QY 623 -----VSLRDQAHA----- 632  
DB 670 CRYCNVLGGEREEARACHATHNACRCRTGFFAHAGFCLEHASCPPGAGVIAPGTPSQN 729  
QY 633 -----PVTLRAS----- 639  
DB 730 TQOCPCPGTFSASSSSSEQPHNCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPG 789  
QY 640 -----FADRTLOWISR-----WG-----QKGLISAPFLPYQP 666

DB 790 ABCECRAVIDFVAFQDISIKRLQRLQLALEAPEGWGPTRAGRAALQLKL-----R 840  
QY 667 RFPEVILLFQEGGLKALNGQGLGATSMNQQAELQRLRLRISG 709  
DB 841 RRLTELLGAGDQALLVRL-----LQALRVARMPG 869

RESULT 6  
US-10-775-180-603  
; Sequence 603, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 603  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-603

Query Match 83.3%; Score 3152.3; DB 17; Length 819;  
Best Local Similarity 76.8%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

QY 1 DAHSEVAHRFKDGBENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHSEVAHRFKDGBENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLPGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLPGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCOADKAACLLP 180  
DB 145 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAFTCCCOADKAACLLP 204  
QY 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLTK 240  
DB 205 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLTK 264  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVENDMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVENDMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEKDVLGMFLYVARHPDYSVVLRLRLAKTVEITTELEK 360

Db 325 DLPSLAADFVSKGVKQVNAEAKDVLGMFLYEAARRHPDYSVVLRLIAKTYETTLEK 384  
QY 361 CAAADPHECYAKVDFDKPLVREPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFDKPLVREPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKEROIKQTALVELVHKHKKPAT 540  
Db 505 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKEROIKQTALVELVHKHKKPAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL-----MSPRLEV 592  
Db 565 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGLATGLTSLDPSLEI 624  
QY 593 -----ATSMNQAL-----EQLREL 706  
Db 625 YKMFVSRREQLLAKNLALQNDIHQQYKILDMVLMKGLFKVLEDSRTVLTAAADVLPDGP 684  
QY 603 SPGQVILVGLVLOPKHFTVSLRQAAHAPVTLRASPADRTLQ-----W- 647  
Db 685 FP-----QDEK-----LKDAFSHV-VENTAFGDDVVLPRPRIVHYVFDHNSNWN 727  
QY 648 -ISRMGQKKLISAPFLFYPPQRFVLELLFQEGGLK--LALNGQGLG----- 690  
Db 728 LLIRWG-----IS-----FCNQTGVFNQPHSPILSLMAQELGISEKDSNFQNPFKI 774  
QY 691 -----ATSMNQAL-----EQLREL 706  
Db 775 DRTEFIPSTDPPQKALREBEKRRKKEKRKEIR 807

RESULT 7

US-10-775-180-605  
; Sequence 605, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 605  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-775-180-605  
Query Match 83.3%; Score 3152.3; DB 17; Length 819;  
Best Local Similarity 76.6%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;  
QY 1 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQOCPPFDHVKLVNVEVTEFAKTCVADESAAE 60  
Db 25 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQOCPPFDHVKLVNVEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECPLOHKDNDPNLPLRVBREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECPLOHKDNDPNLPLRVBREV 144  
QY 121 DVMCCTAFHDNBTFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTTECCOAAADKAAACLLP 180  
Db 145 DVMCCTAFHDNBTFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTTECCOAAADKAAACLLP 204  
QY 181 KLDELDRBEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSQRFPPKAFAEVSCLVTDLT 240  
Db 205 KLDELDRBEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSQRFPPKAFAEVSCLVTDLT 264  
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTTECHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKGVKQVNAEAKDVLGMFLYEAARRHPDYSVVLRLIAKTYETTLEK 360  
Db 325 DLPSLAADFVSKGVKQVNAEAKDVLGMFLYEAARRHPDYSVVLRLIAKTYETTLEK 384  
QY 361 CAAADPHECYAKVDFDKPLVREPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 420  
Db 385 CAAADPHECYAKVDFDKPLVREPQNLIKONCELPOLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKEROIKQTALVELVHKHKKPAT 540  
Db 505 LVNRRPCFSALEVDETYVPKFNATFTFHADICTLSEKEROIKQTALVELVHKHKKPAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL-----MSPRLEV 592  
Db 565 KEQLKAVMDDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGLATGLTSLDPSLEI 624  
QY 593 -----PCSHALPOGL 602  
Db 625 YKMFVSRREQLLAKNLALQNDIHQQYKILDMVLMKGLFKVLEDSRTVLTAAADVLPDGP 684  
QY 603 SPGQVILVGLVLOPKHFTVSLRQAAHAPVTLRASPADRTLQ-----W- 647  
Db 685 FP-----QDEK-----LKDAFSHV-VENTAFGDDVVLPRPRIVHYVFDHNSNWN 727  
QY 648 -ISRMGQKKLISAPFLFYPPQRFVLELLFQEGGLK--LALNGQGLG----- 690  
Db 728 LLIRWG-----IS-----FCNQTGVFNQPHSPILSLMAQELGISEKDSNFQNPFKI 774  
QY 691 -----ATSMNQAL-----EQLREL 706  
Db 775 DRTEFIPSTDPPQKALREBEKRRKKEKRKEIR 807

RESULT 8

US-10-775-204-1614  
; Sequence 1614, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1614  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-1614

Query Match 83.3%; Score 3152.3; DB 18; Length 819;  
Best Local Similarity 76.6%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

QY 1 DAHSEVAHFKDGLGEENFKALVLIAPAFQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 60  
DB 25 DAHSEVAHFKDGLGEENFKALVLIAPAFQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAPHDNEETFLKYLVEIARRHPYFYAPPELLPEAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 DVNCTAPHDNEETFLKYLVEIARRHPYFYAPPELLPEAKRYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELDRDEGKASAKQRLKASQKQFGERAFKAWAVARLSQRFPKAFEAESVSKLVTDLTK 240  
DB 205 KLDELDRDEGKASAKQRLKASQKQFGERAFKAWAVARLSQRFPKAFEAESVSKLVTDLTK 264  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKDVCNVAEAKDVFGLMFLEYARHPDYSVVLLRLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVSKDVCNVAEAKDVFGLMFLEYARHPDYSVVLLRLAKTYETTTLEKC 384  
QY 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNQALLVRYTKVPQVST 420  
DB 385 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNQALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKQRIKKQTALVELVKGKPKAT 540  
DB 505 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKQRIKKQTALVELVKGKPKAT 564  
QY 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEBGKKLVAASQAALGL-----MSPLREV 592

Db 565 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEBGKKLVAASQAALGLATGTLRTSLDLSLEI 624  
QY 593 -----PCSHALPOGL 602  
Db 625 YKKMFVEKEREQLLALKNLQALNDIHQYKILDMVKGLFKVLEDSRTVLTAAADVLPDGP 684  
QY 603 SPGQVIVRGLVLOQPKFTVSLRDOAHAPVTLRASFADRTLO-----W- 647  
Db 685 FP-----ODEK-----LKDAFHSV-VENTAFGDDVVLFRPRIVHYFDHNSNN 727  
QY 648 -ISRWGOKKLIAPFLFYQPRFEVILLFQEGGLK--LALNGOGLG----- 690  
Db 728 LLIRWG-----IS-----FCNQTGVFNQGHPSILSMAQELGISEKDSNFQNPFKI 774  
QY 691 -----ATSMNQAL-----EQLRELR 706  
Db 775 DRTEFIPSTDPFQKALREBEKRRKKEKEKEIR 807

RESULT 9  
US-10-775-204-1616  
; Sequence 1616, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1616  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-1616

Query Match 83.3%; Score 3152.3; DB 18; Length 819;  
Best Local Similarity 76.6%; Pred. No. 2.3e-75;  
Matches 623; Conservative 16; Mismatches 37; Indels 137; Gaps 12;

QY 1 DAHSEVAHFKDGLGEENFKALVLIAPAFQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 60  
DB 25 DAHSEVAHFKDGLGEENFKALVLIAPAFQYLOQCPPEHVKLVNEVTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144  
QY 121 DVNCTAPHDNEETFLKYLVEIARRHPYFYAPPELLPEAKRYKAAFTCCQAADKAAACLLP 180  
DB 145 DVNCTAPHDNEETFLKYLVEIARRHPYFYAPPELLPEAKRYKAAFTCCQAADKAAACLLP 204  
QY 181 KLDELDRDEGKASAKQRLKASQKQFGERAFKAWAVARLSQRFPKAFEAESVSKLVTDLTK 240  
DB 205 KLDELDRDEGKASAKQRLKASQKQFGERAFKAWAVARLSQRFPKAFEAESVSKLVTDLTK 264  
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCHIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKDVCNVAEAKDVFGLMFLEYARHPDYSVVLLRLAKTYETTTLEKC 360  
DB 325 DLPSLAADFVSKDVCNVAEAKDVFGLMFLEYARHPDYSVVLLRLAKTYETTTLEKC 384  
QY 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNQALLVRYTKVPQVST 420  
DB 385 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNQALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKQRIKKQTALVELVKGKPKAT 540  
DB 505 LVNRRPCFSALEVDVETVPKFNATETTFHADICTLSEKQRIKKQTALVELVKGKPKAT 564  
QY 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEBGKKLVAASQAALGL-----MSPLREV 592

```
QY 121 DVMCTAFHNDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHNDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Db 205 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLTk 264
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Db 325 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAMPKPCAEYDLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPKAMPKPCAEYDLSVVLNQLCVLHKTVPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCCKADKCTCFABEGKCLVAASQAALGL-MSPRLEV 592
Db 565 KEQLKAVMDDFAAFVEKCKCCKADKCTCFABEGKCLVAASQAALGLTGLTSLDPSLEI 624
QY 593 -----ATSMNQAL-----EQLREUR 706
Db 625 YKMFVKKRREQLLAKNLALNDIHQYKILDMVLKGLFKVLEDSRTVLTAAVLPDGP 684
QY 603 SPQVILVRLVLPKPHFTVSLDQAAHAPVTLRASADRTLQ-----W- 647
Db 685 FP-----QDEK-----LKDAFHSV-VENTAFGDDVLRPRVHYHYFDHNSWN 727
QY 648 -ISRWGQKLLISAPFLFYPQRFVFFVLLFQEGGLK--LALNQQGLG----- 690
Db 728 LLIRWG---IS-----FCNQTGVNFQPHSPILSLMAQELGISEKDSNFQNPFKI 774
QY 691 -----ATSMNQAL-----EQLREUR 706
Db 775 DRTEPIPTDPFQKALREBEKKRKEKRKEIR 807

RESULT 10
US-10-775-204-1657
; Sequence 1657, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
```

```
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1657
; LENGTH: 794
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-1657
```

```
Query Match 83.3%; Score 3152.2; DB 18; Length 794;
Best Local Similarity 76.3%; Pred. No. 2.2e-75;
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFAQYLOQCPPEHDHVKLVNVEVTEPAKTCVADESAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIATAFAQYLOQCPPEHDHVKLVNVEVTEPAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHNDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHNDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLTk 240
Db 205 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEFAEVSCLVTDLTk 264
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Db 325 DLPSLAADPVESKOVCKNYAEAKOVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEPEQLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAMPKPCAEYDLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPKAMPKPCAEYDLSVVLNQLCVLHKTVPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCCKADKCTCFABEGKCLVAASQAALGLM-SPRLEV 592
Db 565 KEQLKAVMDDFAAFVEKCKCCKADKCTCFABEGKCLVAASQAALGLMTEHSPITPHRRD 624
QY 593 PCSHAL-----PQGLSPG----- 605
Db 625 LCSSRIMLARKIRSDLTALTESYVVKHQGLNKINILNDSADGMVAVASTDQWSELTAEERLQE 684
QY 606 -----QVILVRLVLPKPHFTVSLRD--QAAHAPVTLRASFA----- 641
Db 685 NLQAYRTPHVLAR--LLEDQVQVHFTPTGDFHQAHTHLLQLVAFAFYQIEELMILLEYKI 743
QY 642 -----DRTLQWISRWGQKLLISAPFLFYPQRFVFFVLLLRLLAKTYETTLK 683
```

```
Db 744 PRNEADGMPINVDGGLFEKKL-----WGLKVL-----QE----- 773
Qy 684 LNGQGLGATSMNQALEQLRELRLISGSVQ 712
Db 774 -----LSQWTVRSIHDLRFISSHQ 792

RESULT 11
US-10-775-204-1548
; Sequence 1548, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1548
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-1548

Query Match 83.3%; Score 3152.2; DB 18; Length 809;
Best Local Similarity 76.3%; Pred. No. 2.2e-75;
Matches 617; Conservative 18; Mismatches 36; Indels 138; Gaps 9;

Qy 1 DAHSEVAHRFDLGENFKALVLIAPAYLOQCPEDHVKLVNVEYTFKACTVADSEAE 60
Db 25 DAHSEVAHRFDLGENFKALVLIAPAYLOQCPEDHVKLVNVEYTFKACTVADSEAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPRLVRPEV 144
Qy 121 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 145 DVNCTAFHNDNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFKAEPFVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFKAEPFVSKLVTDLTK 264
Qy 241 VHTECCHGDLLECADRADLAKYICENODSISSKLKECKECPKLLKSHCTICAEVDEMPA 300
Db 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLKECKECPKLLKSHCTICAEVDEMPA 324
```

```
Qy 301 DLPSLAADPVESKDVCKNYAEAKVDPLGMFLYEYARRHPDYSVLWLLLRLLAKTYETLEKC 360
Db 325 DLPSLAADPVESKDVCKNYAEAKVDPLGMFLYEYARRHPDYSVLWLLLRLLAKTYETLEKC 384
Qy 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKQNCLEPEQLGEYKFQNALLVRYTKVPQVST 420
Db 385 CAADAPHECYAKVDFEFPKPLVEEPQNLIKQNCLEPEQLGEYKFQNALLVRYTKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCCGHPKAMPKPCABDYLSVLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCCGHPKAMPKPCABDYLSVLNQLCVLHKEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCPSALEVDETYVPKFNATFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540
Db 505 LVNRRPCPSALEVDETYVPKFNATFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 564
Qy 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGLM-----SPRLVEV 592
Db 565 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGLMFAFTEHSPLTPHRRD 624
Qy 593 PCSHAL-----PQGLSPG----- 605
Db 625 LCSSRSIWLARKIRSOLTALTESYVKHQGLNKNINLDSADGMPVASTDQWSELTEARLQE 684
Qy 606 -----QVIIVRGVLVQEPKHFTVSLD--QAAHAPVTLRASA----- 641
Db 685 NLQAYRTFHVLLAR--LLEDQQVHFTTTEGDFHQAHTLLQVAAFAIYQTEELMILLEYKI 743
Qy 642 -----DRTLQWISRWGQKKLISAPFLFYQRFPEVLLLFQEGGLKLA 683
Db 744 PRNEADGMPINVDGGLFEKKL-----WGLKVL-----QE----- 773
Qy 684 LNGQGLGATSMNQALEQLRELRLISGSVQ 712
Db 774 -----LSQWTVRSIHDLRFISSHQ 792

RESULT 12
US-10-775-204-364
; Sequence 364, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 364
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-364

Query Match      83.3%; Score 3152; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.1e-75;
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
D5 DAHSEVAHRFKDLGEENFKALVLIAPAYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
D5 NCDKSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 180
D5 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 204
QY 145 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 240
D5 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 264
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKAEFAVSKLVTDLTK 240
D5 KLDELDRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKAEFAVSKLVTDLTK 264
QY 241 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENMPA 300
D5 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENMPA 324
QY 301 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 360
D5 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 384
QY 325 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 384
D5 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420
D5 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLLNQLCVLHKEKTPVSDRVTCCCTES 480
D5 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLLNQLCVLHKEKTPVSDRVTCCCTES 504
QY 481 LVNRRPCFSALEVDSTYVPKFNATFTFHADICTLSEKERQIKKQATALVELVKKPKAT 540
D5 LVNRRPCFSALEVDSTYVPKFNATFTFHADICTLSEKERQIKKQATALVELVKKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMSPL----- 590
D5 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMSPL----- 590
QY 591 -----EVPQSHA-----LPGGLS 603
D5 -----EVPQSHA-----LPGGLS 603
QY 625 LLEAKEAQITTCAGHCSLNEQITVPDTKNFYAWKRMVEGQQAQVEVWQGLALLSEAVL 684
D5 LLEAKEAQITTCAGHCSLNEQITVPDTKNFYAWKRMVEGQQAQVEVWQGLALLSEAVL 684
QY 604 PGQVILVRLGLVQEP-----KHFTVSLRDQAHAHPVTLSAFADRTLQWISRW 651
D5 PGQVILVRLGLVQEP-----KHFTVSLRDQAHAHPVTLSAFADRTLQWISRW 651
QY 685 RQGLVLOSSQWPEPLQLHVDKAVGSLSLTLTLR-----ALRA----- 723
D5 RQGLVLOSSQWPEPLQLHVDKAVGSLSLTLTLR-----ALRA----- 723
QY 652 GQKLLISAP-----ELFYPQRFEEVLLLFQEGGLKL 682
D5 GQKLLISAP-----ELFYPQRFEEVLLLFQEGGLKL 682
QY 724 -QKEAISPPDAASAAPLRTITADTF---RKLFRVYSNFLRGKLLK 764
D5 -QKEAISPPDAASAAPLRTITADTF---RKLFRVYSNFLRGKLLK 764
```

## RESULT 13

```
US-10-775-204-366
; Sequence 366, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
```

```
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 366
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-366
```

```
Query Match      83.3%; Score 3152; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.1e-75;
Matches 615; Conservative 10; Mismatches 31; Indels 110; Gaps 7;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
D5 DAHSEVAHRFKDLGEENFKALVLIAPAYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
D5 NCDKSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 180
D5 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 204
QY 145 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 240
D5 DVMCTAFHDNSETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAAACLLP 264
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKAEFAVSKLVTDLTK 240
D5 KLDELDRDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRFPAKAEFAVSKLVTDLTK 264
QY 241 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENMPA 300
D5 VHTCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENMPA 324
QY 301 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 360
D5 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 384
QY 325 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 384
D5 DLPSLAADFVSKQVCKNYAEAKVFLGMFLVEYARRHPDYSVLLLRKLTAKTYETTLKCC 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420
D5 CAAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLLNQLCVLHKEKTPVSDRVTCCCTES 480
D5 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLLNQLCVLHKEKTPVSDRVTCCCTES 504
QY 481 LVNRRPCFSALEVDSTYVPKFNATFTFHADICTLSEKERQIKKQATALVELVKKPKAT 540
D5 LVNRRPCFSALEVDSTYVPKFNATFTFHADICTLSEKERQIKKQATALVELVKKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMSPL----- 590
D5 KEQLKAVMDDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGLMSPL----- 590
```



```
Db 565 KEQLKAVNMDFAAFVEKCKKADDTCTCFABEGKGLVAASQAALGLAPRLICDSRVLERY 624
Qy 591 -----EVPKSHA-----LPOGLS 603
Db 625 LLEAKEARQITTCGAEHCSLNEQITVPTDKVNFYAKWMEVQQAQVAVWQGLLSEAVL 684
Qy 604 PQQVIIVRGLVLEP-----KHFTVSLRDQAHAHAPVTLRASFAADRTLQWISRW 651
Db 685 RQALLVQSSQPWEPLQLHVDKAVSGLSLTLLR-----ALRA-----723
Qy 652 GQKLIISAP-----FLFYPRQFVEVLLLFQEGGLKL 682
Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFRGLKGL 764

RESULT 14
US-10-775-204-245
; Sequence 245, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 245
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-245

Query Match 83.2%; Score 3151; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.2e-75;
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;

Qy 1 DAHKSVAHRFDLGENFKALVLIAPQVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHKSVAHRFDLGENFKALVLIAPQVLOQCPEDHVKLVNEVTEFAKTCVADESAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPV 144
Qy 121 DVNCTAFHNEETFLKKYLYIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLP 180
Db 145 DVNCTAFHNEETFLKKYLYIARRHPYFAPPELLFFAKRYKAAFTTECCQAADKAACLP 204
```

```
Qy 181 KDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPPKAEFAEVSCLVTDLT 240
Db 205 KDELRLDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPPKAEFAEVSCLVTDLT 264
Qy 241 VHTCCGHDLLCECADDRLADLAKYICENODSISLKECCCKPPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCGHDLLCECADDRLADLAKYICENODSISLKECCCKPPLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKQVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKQVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 384
Qy 361 CAADPHCEYAKVDFEPLVEBPONLIKQNCLEFQELGEYKFQVALLVRYTKKVPQVST 420
Db 385 CAADPHCEYAKVDFEPLVEBPONLIKQNCLEFQELGEYKFQVALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDVLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDVLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPKAT 540
Db 505 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPKAT 564
Qy 541 KEQLKAVNMDFAAFVEKCKKADDTCTCFABEGKGLVAASQAALGLMSPL-----590
Db 565 KEQLKAVNMDFAAFVEKCKKADDTCTCFABEGKGLVAASQAALGLAPRLICDSRVLERY 624
Qy 591 -----EVPKSHA-----LPOGLS 603
Db 625 LLEAKEARQITTCGAEHCSLNEITVPTDKVNFYAKWMEVQQAQVAVWQGLLSEAVL 684
Qy 604 PQQVIIVRGLVLEP-----KHFTVSLRDQAHAHAPVTLRASFAADRTLQWISRW 651
Db 685 RQALLVQSSQPWEPLQLHVDKAVSGLSLTLLR-----ALRA-----723
Qy 652 GQKLIISAP-----FLFYPRQFVEVLLLFQEGGLKL 682
Db 724 -QKEAISPPDAASAAPLRTITADTF-----RKLFRVYSNFRGLKGL 764

RESULT 15
US-10-775-204-255
; Sequence 255, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
```

```
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 255
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-255

Query Match      83.2%; Score 3151; DB 18; Length 774;
Best Local Similarity 80.5%; Pred. No. 2.2e-75;
Matches 617; Conservative 9; Mismatches 30; Indels 110; Gaps 8;

Qy 1 DAHSEVAHRPKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 25 DAHSEVAHRPKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEVSKLVTDLTK 264
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVNDENPA 300
Db 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVNDENPA 324
Qy 301 DLPSLAADFVSKDVKCKNYAEAKDVLGMFLYEVARRHPDYSVVLNLLRLAKTYETTLK 360
Db 325 DLPSLAADFVSKDVKCKNYAEAKDVLGMFLYEVARRHPDYSVVLNLLRLAKTYETTLK 384
Qy 361 CAAADPHECYAKVFDEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVFDEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDITYVPKGFNAETFTFHADICTLSEKQIKKQATLALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDITYVPKGFNAETFTFHADICTLSEKQIKKQATLALVELVKHKPKAT 564
Qy 541 KEQLKAVMDDFAAFVEKCKKADDETCFAPEGKLVAAASQAALGMSPR----- 589
Db 565 KEQLKAVMDDFAAFVEKCKKADDETCFAPEGKLVAAASQAALGMSPR----- 624
Qy 590 -LEVPCHALP-----KHFTVSLRDQAHAHPVTLIRASFADRTLQWISRW 603
Db 625 LLEAKEAEAITTGCABHCSLNEAIVTPDKNVNFYAWKMEVQQQAVFVWQGLALLSEAVL 684
Qy 604 PQQVIIVRGLVLEP-----KHFTVSLRDQAHAHPVTLIRASFADRTLQWISRW 651
Db 685 RQALLVASSQWEPQLQHLVDKAVSGLSLTLTLR-----ALRA----- 723
Qy 652 GQKLIASAP-----FLVPQRPFFEVLLIFQEGGLKL 682
Db 724 -QKEAISPPDAASAAPLRTITADTF---RKLFRVYSNFLRGKLL 764

RESULT 16
US-10-775-204-256
; Sequence 256, Application US/10775204
; Publication No. US2005018664A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 256
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-256

Query Match      83.2%; Score 3151; DB 18; Length 774;
Best Local Similarity 80.3%; Pred. No. 2.2e-75;
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;

Qy 1 DAHSEVAHRPKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 25 DAHSEVAHRPKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFPAKAEVSKLVTDLTK 264
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVNDENPA 300
Db 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVNDENPA 324
Qy 301 DLPSLAADFVSKDVKCKNYAEAKDVLGMFLYEVARRHPDYSVVLNLLRLAKTYETTLK 360
Db 325 DLPSLAADFVSKDVKCKNYAEAKDVLGMFLYEVARRHPDYSVVLNLLRLAKTYETTLK 384
Qy 361 CAAADPHECYAKVFDEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVFDEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 444
Qy 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
```



;; PRIOR FILING DATE: 2002-10-11  
;; PRIOR APPLICATION NUMBER: 60/420,246  
;; PRIOR FILING DATE: 2002-10-23  
;; PRIOR APPLICATION NUMBER: 60/423,623  
;; PRIOR FILING DATE: 2002-11-05  
;; PRIOR APPLICATION NUMBER: 60/351,360  
;; PRIOR FILING DATE: 2002-01-28  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 222  
;; SOFTWARE: Patentin Ver. 2.0  
;; SEQ ID NO 302  
;; LENGTH: 774  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-775-204-302

Query Match 83.2%; Score 3151; DB 18; Length 774;  
Best Local Similarity 80.3%; Pred. No. 2.2e-75;  
Matches 615; Conservative 9; Mismatches 32; Indels 110; Gaps 7;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 60  
Db 25 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204

Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTLDTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTLDTK 264

Qy 241 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360  
Db 325 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 384

Qy 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 420  
Db 385 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 444

Qy 421 PTLVEVSRLGKVGSKCKKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480  
Db 445 PTLVEVSRLGKVGSKCKKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTCKCTES 504

Qy 481 LVNRRPCFSALEVDYTPKPFNAETFTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540  
Db 505 LVNRRPCFSALEVDYTPKPFNAETFTFHADICTLSEKERQIKKOTALVELVXHKPKAT 564

Qy 541 KEQLKAVMDDFAAFVEKCKADDKETCPAEBGKLVAAASQAALGMSRPL----- 590  
Db 565 KEQLKAVMDDFAAFVEKCKADDKETCPAEBGKLVAAASQAALGMSRPL----- 624

Qy 591 -----EVPKSHA-----LPGQLS 603  
Db 625 LLEAKEAENITTGCAEHCSLNEINITVPTDKVNFYAWKMEVGGQAVQVWQGLALLSEAVL 684

Qy 604 PGQVLIIVRGLVQEP-----KHFTVSLRDQAHAHPVTLRAFADRTLQWISRW 651  
Db 685 RGQALVNSQWEPFLQLHVDKAVSGLSLTLTLR-----ALRA----- 723

Qy 652 GQKKLISAP-----FLFYPQRFVEVLLIFQEGGLKL 682  
Db 724 -QKEAISPPDAASAAPLRTITADTE- - - - -RKLFRVYSNFLRGKLL 764

RESULT 19  
US-10-775-180-72  
;; Sequence 72, Application US/10775180  
;; Publication No. US20050054570A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen, Craig A.  
;; TITLE OF INVENTION: Albumin Fusion Proteins  
;; FILE REFERENCE: PFS74  
;; CURRENT FILING DATE: 2004-02-11  
;; PRIOR APPLICATION NUMBER: PCT/US02/40892  
;; PRIOR FILING DATE: 2002-12-23  
;; PRIOR APPLICATION NUMBER: 60/341,811  
;; PRIOR FILING DATE: 2001-12-21  
;; PRIOR APPLICATION NUMBER: 60/360,000  
;; PRIOR FILING DATE: 2002-02-28  
;; PRIOR APPLICATION NUMBER: 60/378,950  
;; PRIOR FILING DATE: 2002-05-10  
;; PRIOR APPLICATION NUMBER: 60/398,008  
;; PRIOR FILING DATE: 2002-07-24  
;; PRIOR APPLICATION NUMBER: 60/411,355  
;; PRIOR FILING DATE: 2002-09-18  
;; PRIOR APPLICATION NUMBER: 60/414,984  
;; PRIOR FILING DATE: 2002-10-02  
;; PRIOR APPLICATION NUMBER: 60/417,611  
;; PRIOR FILING DATE: 2002-10-11  
;; PRIOR APPLICATION NUMBER: 60/420,246  
;; PRIOR FILING DATE: 2002-10-23  
;; PRIOR APPLICATION NUMBER: 60/423,623  
;; PRIOR FILING DATE: 2002-11-05  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 858  
;; SOFTWARE: Patentin Ver. 2.0  
;; SEQ ID NO 72  
;; LENGTH: 868  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-775-180-72

Query Match 83.2%; Score 3151; DB 17; Length 868;  
Best Local Similarity 70.5%; Pred. No. 2.7e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 60  
Db 25 DAHSEVAHRPKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNVEVTEFAKTCVADSSAE 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
Db 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 204

Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTLDTK 240  
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRFPPKAEFAEVSCLVTLDTK 264

Qy 241 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 360  
Db 325 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLK 384

Qy 361 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 420  
Db 385 CAADPHCEYAKVDFEFPKPLVEEPONLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 444

Qy 421 PTLVEVSRLGKVGSKCKKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480

Db 445 PTLVEVSRLGKVGSKCKHPKAEKMPKCAEDYLSVNLQCLVHLKHTFVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKQRIKKQTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKQRIKKQTALVELVGHKPKAT 564  
Qy 541 KEQLKAVNMDFAAFVEKCKKADDETCFAEKGKLVAAASQAALGLMSPLRLEVP 593  
Db 565 KEQLKAVNMDFAAFVEKCKKADDETCFAEKGKLVAAASQAALGLVA---ETPTYPWRDA 621  
Qy 594 -----CSHALPQGLSPGVIIVRGLVLQE-----PKHFT----- 622  
Db 622 ETGERLVCAQCQP-----GTFVQPCRRDSPTTCGCPFRHVTQFWNYLRCR 669  
Qy 623 -----VSLRDQAHA----- 632  
Db 670 YCNVLCGEREBEARACHATHNRACRCRTGFFAHAGFCLHASCPCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 COPCPGTFSSSSSSSEQCPHRNCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPGAE 789  
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQRLQALBAPEGWGPTPRAGRAALQLKL-----RRR 840  
Qy 669 FEVLLFQEGGLKALNGQGLGATSMNQALBQLRELRIISG 709  
Db 841 LTELGAQDQALLVRL-----LQALRVARMGP 867

RESULT 20

US-10-775-204-219  
; Sequence 219, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 222  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 219  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-204-219

Query Match 83.2%; Score 3151; DB 18; Length 868;  
Best Local Similarity 70.5%; Pred. No. 2.7e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;  
Qy 1 DAHSEVAHFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 60  
Db 25 DAHSEVAHFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNVEVTEFAKTCVADESAAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREV 144  
Qy 121 DVMCTAFHDNEETFLKYLVIARRHPYFVAPPELLFFAKRYKAAAFTECCOAAADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLKYLVIARRHPYFVAPPELLFFAKRYKAAAFTECCOAAADKAACLLP 204  
Qy 181 KLDELREGKASSAKORLKASLOKFGERAFAKAVARLSORFPKAEFAEVSCLVDTLTK 240  
Db 205 KLDELREGKASSAKORLKASLOKFGERAFAKAVARLSORFPKAEFAEVSCLVDTLTK 264  
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYEVARRHPDYSVULLLHLAKTYETTLK 360  
Db 325 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYEVARRHPDYSVULLLHLAKTYETTLK 384  
Qy 361 CAADPHCEYAKVDFDEKPLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKVPQVST 420  
Db 385 CAADPHCEYAKVDFDEKPLVEEPQNLIKONCELFEOQGEYKFNALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRLGKVGSKCKHPKAEKMPKCAEDYLSVNLQCLVHLKHTFVSDRVTKCCTES 480  
Db 445 PTLVEVSRLGKVGSKCKHPKAEKMPKCAEDYLSVNLQCLVHLKHTFVSDRVTKCCTES 504  
Qy 481 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKQRIKKQTALVELVGHKPKAT 540  
Db 505 LVNRRPCFSALEVDYVYVPEKFNATFTFFHADICTLSEKQRIKKQTALVELVGHKPKAT 564  
Qy 541 KEQLKAVNMDFAAFVEKCKKADDETCFAEKGKLVAAASQAALGLMSPLRLEVP----- 593  
Db 565 KEQLKAVNMDFAAFVEKCKKADDETCFAEKGKLVAAASQAALGLVA---ETPTYPWRDA 621  
Qy 594 -----CSHALPQGLSPGVIIVRGLVLQE-----PKHFT----- 622  
Db 622 ETGERLVCAQCQP-----GTFVQPCRRDSPTTCGCPFRHVTQFWNYLRCR 669  
Qy 623 -----VSLRDQAHA----- 632  
Db 670 YCNVLCGEREBEARACHATHNRACRCRTGFFAHAGFCLHASCPCPGAGVIAPGTPSQNTQ 729  
Qy 633 -----PVTLRAS----- 639  
Db 730 COPCPGTFSSSSSSSEQCPHRNCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPGAE 789  
Qy 640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYPORF 668  
Db 790 ECERAVIDFVAFQDISIKRLQRLQALBAPEGWGPTPRAGRAALQLKL-----RRR 840  
Qy 669 FEVLLFQEGGLKALNGQGLGATSMNQALBQLRELRIISG 709  
Db 841 LTELGAQDQALLVRL-----LQALRVARMGP 867

RESULT 21

US-10-775-180-71  
; Sequence 71, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.

```

; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 71
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-775-180-71

Query Match      83.2%; Score 3151; DB 17; Length 874;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRRFKDGEENFKALVLIAPFAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 25 DAHSEVAHRRFKDGEENFKALVLIAPFAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSSAE 84

Qy 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVREPV 120
Db 85 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVREPV 144

Qy 121 DVMCTAFHDNBEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 145 DVMCTAFHDNBEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 204

Qy 181 KLDELRDGKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
Db 205 KLDELRDGKASSAKORLKCASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 264

Qy 241 VHTCECHGDLLECCADDRADLAKYICENQDSISLKLKECCXPLEKSHCIAEVENDEMPA 300
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISLKLKECCXPLEKSHCIAEVENDEMPA 324

Qy 301 DLPSLAADFVSKDCKNKAABKOVFLQWFLYFYARRHPDYSVLLRLAKTYETFTLEKC 360
Db 325 DLPSLAADFVSKDCKNKAABKOVFLQWFLYFYARRHPDYSVLLRLAKTYETFTLEKC 384

Qy 361 CAADDPHECYAKVDFEKPFLVEEPONLJKONCELPQOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAADDPHECYAKVDFEKPFLVEEPONLJKONCELPQOLGEYKFNALLVRYTKKVPQVST 444

Qy 421 PTLVEVSRLNGKSGKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNGKSGKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504

Qy 481 LVNRRPCFSALEVDETYVPKGFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPAT 540
Db 505 LVNRRPCFSALEVDETYVPKGFNAETFTFHADICTLSEKERQIKQTALVELVKKHKPAT 564

; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 71
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-775-180-71

RESULT 22
US-10-775-204-218
; Sequence 218, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haeeltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 218
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-775-204-218

Query Match      83.2%; Score 3151; DB 18; Length 874;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRRFKDGEENFKALVLIAPFAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 11 DAHSEVAHRRFKDGEENFKALVLIAPFAQYLQOCPPFEDHVKLVNEVTEFAKTCVADSSAE 60
```

Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 144  
Qy 121 DVMCTAFHNEBTFKLYLIEIARHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180  
Db 145 DVMCTAFHNEBTFKLYLIEIARHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264  
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKOVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKOVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETTLK 384  
Qy 361 CAADAPHECYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFONALLVRYTKVPQVST 420  
Db 385 CAADAPHECYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFONALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
Qy 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKQOTALVELVKGHPKAT 540  
Db 505 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKQOTALVELVKGHPKAT 564  
Qy 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGLVA-----ETPTVWRDA 593  
Db 565 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGLVA-----ETPTVWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIIVRGLVQE-----PKHFT----- 622  
Db 622 ETGERLVCAQCPP-----GTFVQRCRRDSPTTCGCPFRHVTQFWNYLERCR 669

623 -----VSLRDOAHA----- 632  
670 YCNVLGGEREEBARACHATHNACRCRTGFFAHAGFLEHASCPPGAGVIAPGTFPSQNTQ 729  
633 -----PVTLRAS----- 639  
730 CQCPPEGTFSSSSSEQCQPHRNCTALGLALNVPGSSSHDTLCTSGTFPLSTRVPGAE 789  
640 -----FADRTLQWISR-----WG-----QKKLISAPFLFYPQRF 668  
790 EGERAVIDFVAQDISIKELQRLQLALEAPEGGPTPRAGRAALQLKL-----RRR 840  
669 FEVLLLFQEGGLKIALNGOGLGATSNMNOALSOQLRELISG 709  
841 LTELGAQDQGALLVRL-----LOALRVARMFG 867

RESULT 23  
US-10-775-180-70  
; Sequence 70, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10/775,180  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811

; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 70  
; LENGTH: 880  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-70

Query Match 83.2%; Score 3151; DB 17; Length 880;  
Best Local Similarity 70.5%; Pred. No. 2.8e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60  
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREV 144  
Qy 121 DVMCTAFHNEBTFKLYLIEIARHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 180  
Db 145 DVMCTAFHNEBTFKLYLIEIARHPYFAPPELLFFAKRYKAAATECCQAADKAACLLP 204  
Qy 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
Db 205 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264  
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 324  
Qy 301 DLPSLAADFVESKOVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETTLK 360  
Db 325 DLPSLAADFVESKOVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETTLK 384  
Qy 361 CAADAPHECYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFONALLVRYTKVPQVST 420  
Db 385 CAADAPHECYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFONALLVRYTKVPQVST 444  
Qy 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCTES 504  
Qy 481 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKQOTALVELVKGHPKAT 540  
Db 505 LVNRRPCFSALEVDITYVPKEFNAETFTPHADICTLSEKERQIKQOTALVELVKGHPKAT 564  
Qy 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGLVA-----ETPTVWRDA 593  
Db 565 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGLVA-----ETPTVWRDA 621  
Qy 594 -----CSHALPQGLSPQGVIIIVRGLVQE-----PKHFT----- 622  
Db 622 ETGERLVCAQCPP-----GTFVQRCRRDSPTTCGCPFRHVTQFWNYLERCR 669

```
QY 623 -----VSLRDOAHNA----- 632
Db 670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGPSQNTQ 729
QY 633 -----PVTLRAS----- 639
Db 730 QCPCPGTFSSSSSSSQCPHRNCTALGLALNVPGSSSHDTLCTSGPPLSTRVPGAE 789
QY 640 -----FADRTLQWISR-----WG-----OKKLIISAPFLFYPPORF 668
Db 790 ECERAVIDFVAFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
QY 669 FEVLLPFOEGGLKALNGQGLGATSMNQQALEQLRELRIISG 709
Db 841 LTELLGAQDQDGLLVRL-----LQALRVARMGP 867

RESULT 24
US-10-775-180-74
; Sequence 74, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 74
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-74

Query Match 83.2%; Score 3151; DB 17; Length 880;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIJAFQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIJAFQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECLOHKDNDPNLPRLVREPV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECLOHKDNDPNLPRLVREPV 144
QY 121 DVMCTAFHDNEFTFLKKYLYETARRHPYFAPPELLFFAKRYKAAFTCECQAADKAACLLP 180
Db 145 DVMCTAFHDNEFTFLKKYLYETARRHPYFAPPELLFFAKRYKAAFTCECQAADKAACLLP 204
```

```
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSORFPKAEFAVSKLVDLTIK 240
Db 205 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWAVARLSORFPKAEFAVSKLVDLTIK 264
QY 241 VHTECHGDILLECADDRAIDLAKYICENQDISISKLKECCCKPPLLEKSHCIAEVENDEMPA 300
Db 265 VHTECHGDILLECADDRAIDLAKYICENQDISISKLKECCCKPPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLLRLLAKTVEITTEK 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLLRLLAKTVEITTEK 384
QY 361 CAAADPHECYAKVDFEPLVEBPONLIKONCELFEOLOGYKFTQNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEPLVEBPONLIKONCELFEOLOGYKFTQNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKEPAT 540
Db 505 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKEPAT 564
QY 541 KEQLKAVMDDFAAFVEKCCKADDKETCFABEGKKLVAASQAALGLMSPLREVP----- 593
Db 565 KEQLKAVMDDFAAFVEKCCKADDKETCFABEGKKLVAASQAALGLVA---ETPTYWPRDA 621
QY 594 -----CSHALPQGLSPGGVIVIRGLVLOE-----PKHFT----- 622
Db 622 ETGERLVCAQCPP-----GTFVQRPCKRDSPTTCGCPFRHYTQFMNLYRRCR 669
QY 623 -----VSLRDOAHNA----- 632
Db 670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGPSQNTQ 729
QY 633 -----PVTLRAS----- 639
Db 730 QCPCPGTFSSSSSSSQCPHRNCTALGLALNVPGSSSHDTLCTSGPPLSTRVPGAE 789
QY 640 -----FADRTLQWISR-----WG-----OKKLIISAPFLFYPPORF 668
Db 790 ECERAVIDFVAFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
QY 669 FEVLLPFOEGGLKALNGQGLGATSMNQQALEQLRELRIISG 709
Db 841 LTELLGAQDQDGLLVRL-----LQALRVARMGP 867

RESULT 25
US-10-775-180-79
; Sequence 79, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF574
; CURRENT FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
```



QY 640 -----FADRTLOWISR-----WG-----QKKLISAPFLVFPORF 668  
DB 790 ECERAVIDFAFQDISIKRQLQRLQALBAPEGWGTTPRAGRAALQKL-----RRR 840  
QY 669 FEVILLFOEGGLKALNGQGLGATSMNQALQELRELRI 709  
DB 841 LTELGAQDQALLVRL-----LQALVARMPG 867

RESULT 26  
US-10-775-180-87  
; Sequence 87, Application US/10775180  
; Publication No. US20050054570A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF574  
; CURRENT APPLICATION NUMBER: US/10775,180  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: PCT/US02/40892  
; PRIOR FILING DATE: 2002-12-23  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 858  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 87  
; LENGTH: 880  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-180-87

Query Match 83.2%; Score 3151; DB 17; Length 880;  
Best Local Similarity 70.5%; Pred. No. 2.8e-75;  
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHSEVAHRFDLGEENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADES 60  
DB 25 DAHSEVAHRFDLGEENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADES 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144  
QY 121 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264  
QY 241 VHTCCGHGDLLECADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCGHGDLLECADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 324

QY 1 DAHSEVAHRFDLGEENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADES 60  
DB 25 DAHSEVAHRFDLGEENFKALVLAFAQYLOQCPEDHVKLVNEVTEFAKTCVADES 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 144  
QY 121 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHNDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264  
QY 241 VHTCCGHGDLLECADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCGHGDLLECADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVSKDVCNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETTTLEKC 360  
DB 325 DLPSLAADFVSKDVCNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETTTLEKC 384  
QY 361 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420  
DB 385 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGVSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGVSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRFCFSALEVDETYVPKEFNAETFTTHADICTLSEKERQIKKOTALVELVGHKPKAT 540  
DB 505 LVNRRFCFSALEVDETYVPKEFNAETFTTHADICTLSEKERQIKKOTALVELVGHKPKAT 564  
QY 541 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAQALGLMSPLRLEVP----- 593  
DB 565 KEQLKAVMDDFAAFVEKCKKADDDKTCFAEKGKLVAAQALGLVA---ETPTYWRDA 621  
QY 594 -----CSHALPQGLSPGVIIIVRGLVQE-----PKHFT----- 622  
DB 622 ETGERLVCAQCP-----GTFVQRPCRDSPTTCGCPPRHRYTFQWNYLERCR 669  
QY 623 -----VSLRQAAHA----- 632  
DB 670 YCNVLGEREEBARACHATHNRACRCKTGPFHAHAGFLEHASCPPGAGVIAPGTPSQNTQ 729  
QY 633 -----PVTLRAS----- 639  
DB 730 CQCPPTGTFSSASSSSSEQCPHQRCTALGLALNVPGSSSHDTLCTSGTGPPLSTRVPGAE 789

```
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYVEARRHPDYSVLLLELAKTYETTLK 360
D 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYVEARRHPDYSVLLLELAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEKPFLVEEONLQKCELFQOLGEYKFNQALVRYTKKVPQVST 420
D 385 CAADPHCEYAKVDFEKPFLVEEONLQKCELFQOLGEYKFNQALVRYTKKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
D 445 PTLVEVSRLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540
D 505 LVNRRPCFSALEVDETYVPKEFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDETCFAEBGKKLVAASQAALGLMSPRLEVP 593
D 565 KEQLKAVMDDFAAFVEKCKKADDETCFAEBGKKLVAASQAALGLVA---ETPTYPWRDA 621
QY 594 -----CSHALPOGLSPGVIIIRGLVLQE-----PKHFT----- 622
D 622 ETGERLVCAQCQP-----GTFVQPCRRDSTTCGCPPRHYTQFWNYLRCR 669
QY 623 -----VSLRDQAHA----- 632
D 670 YCNVLCGEREBEARACHATHNRACRGTGFFAHAGFCLHEASCPPGAGVIAPGTPSQNTQ 729
QY 633 -----PVTLRAS----- 639
D 730 CQPCPPGTFSASSSSSECOQPHRNTALGLALNVPGSSSHDTLCTSCGFPFLSTRVPGAE 789
QY 640 -----FADRTLQWISR-----WG-----QKKLISAPFLVYPORF 668
D 790 ECERAVIDFVAFQDISIKRLQRLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
QY 669 FEVLLFQEGGLKALNQGLGATSMNQALEQLRELISG 709
D 841 LTELGAQDQALLVRL-----LQALVARMPG 867
```

## RESULT 27

```
US-10-775-180-90
; Sequence 90, Application US/10775180
; Publication No. US20050054570A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS74
; CURRENT APPLICATION NUMBER: US/10/775,180
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: PCT/US02/40892
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 858
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 90
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-180-90
```

```
Query Match      83.2%; Score 3151; DB 17; Length 880;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;
```

```
QY 1 DAHSEVAHRPKDILGEENFKALVLIAPQYLLQCCPFEDHVKLVNVEVTEFAKTCVADESAE 60
D 25 DAHSEVAHRPKDILGEENFKALVLIAPQYLLQCCPFEDHVKLVNVEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLFTTYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
D 85 NCDKSLHTLFGDKLCTVATLFTTYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNBEETFLKKLYBIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
D 145 DVMCTAFHDNBEETFLKKLYBIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 204
QY 181 KLDELIRDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEFAEYKLVTLDTK 240
D 205 KLDELIRDEGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEFAEYKLVTLDTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
D 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYVEARRHPDYSVLLLELAKTYETTLK 360
D 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYVEARRHPDYSVLLLELAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEKPFLVEEONLQKCELFQOLGEYKFNQALVRYTKKVPQVST 420
D 385 CAADPHCEYAKVDFEKPFLVEEONLQKCELFQOLGEYKFNQALVRYTKKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
D 445 PTLVEVSRLGKVGSKCCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540
D 505 LVNRRPCFSALEVDETYVPKEFNATFTFHADICTLSEKERQIKKQATLVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDETCFAEBGKKLVAASQAALGLMSPRLEVP 593
D 565 KEQLKAVMDDFAAFVEKCKKADDETCFAEBGKKLVAASQAALGLVA---ETPTYPWRDA 621
QY 594 -----CSHALPOGLSPGVIIIRGLVLQE-----PKHFT----- 622
D 622 ETGERLVCAQCQP-----GTFVQPCRRDSTTCGCPPRHYTQFWNYLRCR 669
QY 623 -----VSLRDQAHA----- 632
D 670 YCNVLCGEREBEARACHATHNRACRGTGFFAHAGFCLHEASCPPGAGVIAPGTPSQNTQ 729
QY 633 -----PVTLRAS----- 639
D 730 CQPCPPGTFSASSSSSECOQPHRNTALGLALNVPGSSSHDTLCTSCGFPFLSTRVPGAE 789
QY 640 -----FADRTLQWISR-----WG-----QKKLISAPFLVYPORF 668
D 790 ECERAVIDFVAFQDISIKRLQRLQALEAPEGWGPTPRAGRAALQLKL-----RRR 840
QY 669 FEVLLFQEGGLKALNQGLGATSMNQALEQLRELISG 709
D 841 LTELGAQDQALLVRL-----LQALVARMPG 867
```

```

RESULT 28
; Sequence 217, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-217

Query Match      83.2%; Score 3151; DB 18; Length 880;
Best Local Similarity 70.5%; Pred. No. 2.8e-75;
Matches 621; Conservative 17; Mismatches 33; Indels 210; Gaps 11;

QY 1 DAHKEVAHRFDKLGNEENFKALVLTAFAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHKEVAHRFDKLGNEENFKALVLTAFAYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOKHDDNPNLPLRVPRV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOKHDDNPNLPLRVPRV 144

QY 121 DVMCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCECCQAADKAACLLP 180
DB 145 DVMCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCECCQAADKAACLLP 204

QY 181 KLDELDEGKASAKQRLKASLOKGERAFKAWAVARLUSQRPKAEVSKLVTDLTK 240
DB 205 KLDELDEGKASAKQRLKASLOKGERAFKAWAVARLUSQRPKAEVSKLVTDLTK 264

QY 241 VHECCHGDLLECCADRADLAKYICENQDISISKLKECCKEPKLLEKSHCIAEVENDMPA 300
DB 265 VHECCHGDLLECCADRADLAKYICENQDISISKLKECCKEPKLLEKSHCIAEVENDMPA 324

QY 301 DLPSLAADFVESKDVCKNVAEKDVLGMLFLEYEARHPDYSVLLLRLLAKTYETTTLEK 360
DB 325 DLPSLAADFVESKDVCKNVAEKDVLGMLFLEYEARHPDYSVLLLRLLAKTYETTTLEK 384

QY 361 CAAADPHECVAKVDFBPKLVBEPPQNLIKQNCFLFQGLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAAADPHECVAKVDFBPKLVBEPPQNLIKQNCFLFQGLGEYKFNALLVRYTKKVPQVST 444

421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
445 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
481 LVNRRPCFSALEVEDETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKHKKPKAT 540
505 LVNRRPCFSALEVEDETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKHKKPKAT 564
541 KEQLKAVMDDPAAFEVKCKKADDEKTCFAEEGKKLVAAASQAALGMSPLREVP----- 593
565 KEQLKAVMDDPAAFEVKCKKADDEKTCFAEEGKKLVAAASQAALGLVA---ETPTYPWRDA 621
594 -----CSHALPOGLSPGVIIVRGLVLOB-----PKHFT----- 622
622 ETGERLVCAQCPP-----GTFVQRPCCRRDSTTCGCPCPRHRYTOFWNYLRCR 669
623 -----VSLRDOAANA----- 632
670 YCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGTSPQNTQ 729
633 -----PVTLRAS----- 639
730 CQCPPGTGFSSSSSSSEQCPHRNCTALGLALNVPGSSSHDTLCTCTGTFPLSTRVPGAE 789
640 -----PADRTLOWISR-----WG-----OKKLISAPFLFYPQRF 668
790 ECERAVIDFVAFQDISIKRLQRLQALBAPEGWGTTPRAGRAALQLKL-----RRR 840
669 FEVLLIFQEGGLKALNGQGLGATSMNQALQEQLRELRISSG 709
841 LTELIGAQDGALLVRL-----LQALVARMPG 867

RESULT 29
US-10-775-204-221
; Sequence 221, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: Patentin Ver. 2.0
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-217
```





**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 44 Seconds  
(without alignments)  
1570.083 Million cell updates/sec

Title: US-10-933-523-18  
Perfect score: 3785  
Sequence: 1 DAHSEVAHRPKDLGEENFK.....leqireirsgsvqlycvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	82.0	609	1 ABHUS	serum albumin prec
2	2942	77.7	600	2 A47391	serum albumin prec
3	2820	69.2	608	2 S57632	serum albumin prec
4	2475.9	65.4	607	1 ABHOS	serum albumin prec
5	2446.9	64.6	607	1 ABHOS	serum albumin prec
6	2432.9	64.3	607	1 ABHOS	serum albumin prec
7	2426	64.1	608	1 ABHOS	serum albumin prec
8	2416.8	63.9	605	1 ABPGS	serum albumin prec
9	2387	63.1	609	2 JCS838	albumin - Mongolia
10	1861	49.2	453	2 A05139	serum albumin - mo
11	1557.9	41.2	615	1 ABCHS	serum albumin prec
12	1256.7	33.2	609	2 JCS4258	alpha-fetoprotein
13	1252.7	33.1	609	1 FPHU	alpha-fetoprotein
14	1245.7	32.9	609	1 FPGO	alpha-fetoprotein
15	1206.6	31.9	607	1 ABXL72	74K albumin prec
16	1181.9	31.2	265	2 I46986	albumin - dog (fra
17	1178.3	31.1	608	1 ABXL68	68K serum albumin
18	1086.8	28.7	605	1 FPM	alpha-fetoprotein
19	1071.6	28.3	611	1 FPPT	alpha-fetoprotein
20	1059	28.0	599	1 A54906	afamin precursor -
21	933.4	24.7	614	2 S59517	serum albumin prec
22	932	24.6	608	2 A53195	afamin precursor
23	765.3	20.2	608	1 ABONS1	serum albumin 1 pr
24	761.3	20.1	608	1 ABONS2	serum albumin 2 pr
25	700.4	18.5	382	2 A37253	serum albumin - bu
26	531.4	14.0	1423	1 S27941	serum albumin - se
27	421	11.1	476	1 VYRTD	vitamin D-binding
28	417.5	11.0	474	1 VYHUD	vitamin D-binding
29	414	10.9	472	1 A35327	vitamin D-binding

30	332.1	8.8	3796	2	T18514	lysosomal traffic
31	329.6	8.7	1927	2	G64585	cag pathogenicity
32	327.2	8.6	1819	2	A71928	cag island protein
33	324	8.6	2154	2	A84669	hypothetical prote
34	322.8	8.5	2429	1	SJHUA	spectrin alpha cha
35	321.6	8.5	3951	1	VFIHB1	F1 protein - avian
36	320.5	8.5	1920	2	A53188	paricetrin - mous
37	318.2	8.4	3006	2	T28625	variant-specific s
38	317.6	8.4	2210	1	RRXPTV	genome polyprotein
39	317.2	8.4	2168	2	T30171	genome polyprotein
40	316	8.3	4872	2	S27272	ryanodine receptor
41	313.4	8.3	4063	2	T42993	probable spectrin
42	313.4	8.3	4101	2	T23630	hypothetical prote
43	312.2	8.2	26926	1	I38344	titin, cardiac mus
44	311.7	8.2	4868	2	B54161	ryanodine-binding
45	310	8.2	3051	2	S42373	hypothetical prote
46	309.5	8.2	1303	2	E96805	polyketide synthas
47	309.5	8.2	4273	2	C69679	hypothetical prote
48	308.4	8.1	1529	2	T20986	nuclear migration
49	308.4	8.1	2748	2	S57976	lysosomal traffic
50	307.5	8.1	3788	2	T30851	ataxia telangiecta
51	307.4	8.1	2514	2	T37320	ryanodine receptor
52	306.8	8.1	4969	2	A37113	nuclear/mitotic ap
53	306.5	8.1	2253	2	T30336	hypothetical prote
54	306.5	8.1	2619	2	T24588	guanine nucleotide
55	306.2	8.1	2488	2	T42739	hemocytin - silkw
56	305.9	8.1	3133	2	S52093	calcium-binding pr
57	304.6	8.0	1560	2	T30282	elastic titin - hu
58	304.6	8.0	7962	2	I38346	genome polyprotein
59	304.1	8.0	2210	1	RRXPIC	protein T6D22.14 [
60	304	8.0	2254	2	D86215	cation-independent
61	303	8.0	2470	2	I50726	neurofibromatosis-
62	302	8.0	2818	2	B55282	neurofibromin I -
63	301.4	8.0	2820	2	JCS196	polyketide synthas
64	300.1	7.9	4447	2	A69679	type I fatty acid
65	300.1	7.9	8243	2	T31307	major merozoite su
66	299.7	7.9	1772	2	A45532	telomerase-associ
67	299.7	7.9	2629	2	T30987	protein kinase f5c
68	298.4	7.9	4128	2	JCS306	mprp protein - mo
69	298.1	7.9	1979	2	JW0059	dynein beta heavy
70	298.1	7.9	4466	1	S17231	giantin - human
71	297.8	7.9	3225	2	I52300	giantin - human
72	297.8	7.9	3259	1	A56539	insulin-like growt
73	297.5	7.9	2491	1	A28372	hypothetical prote
74	297.3	7.9	3672	2	T23433	probable laminin a
75	297.3	7.9	3704	2	T37316	bacitracin synthet
76	297.2	7.9	2607	2	T31678	genome polyprotein
77	297.1	7.8	3068	1	A44062	hypothetical prote
78	296.9	7.8	1896	2	T01490	genome polyprotein
79	296.8	7.8	2269	1	JQ1750	polyketide synthas
80	296.8	7.8	2478	2	AH2140	364K Golgi complex
81	296.7	7.8	3187	2	JCS837	variant-specific s
82	296.5	7.8	2706	2	T28155	variant-specific s
83	296.4	7.8	1932	2	T25525	variant-specific s
84	296.3	7.8	1711	2	C71625	breast cancer tumo
85	296.3	7.8	3418	1	G02334	breast cancer susc
86	296.1	7.8	3329	2	T42205	breast cancer tumo
87	296.1	7.8	3329	2	T30904	myosin heavy chain
88	295.9	7.8	1515	2	S51824	spectrin alpha cha
89	295.9	7.8	1640	2	T09522	clathrin heavy cha
90	295.6	7.8	1640	2	T09522	beige protein homo
91	295.2	7.8	3788	2	T13960	gramicidin S synth
92	295	7.8	4450	2	JX0340	variant surface pr
93	294.8	7.8	3026	2	T28431	protein F21H11.2 [
94	294.7	7.8	2700	2	D88450	neurofibromin - mo
95	294.7	7.8	2825	2	I54352	utrophin - human
96	294.6	7.8	3433	1	S28381	hypothetical prote
97	294.5	7.8	1780	2	T17272	genome polyprotein
98	294.5	7.8	2105	1	A44059	genome polyprotein
99	294.5	7.8	4488	1	RRIHM2	protein P2K11.14 [
100	294.4	7.8	1518	2	D96660	

Fused





R;He, X.M.; Carter, D.C.  
 Nature 358, 209-215, 1992  
 A>Title: Atomic structure and chemistry of human serum albumin.  
 A;Reference number: A46756; MUID:92334427; PMID:1630489  
 A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
 R;Brown, J.R.; Shockley, P.; Behrens, P.Q.  
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,  
 A;Reference number: A94442  
 A;Contents: annotation; three-dimensional structure and disulfide bonds  
 R;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.  
 Collect. Czech. Chem. Commun. 42, 564-579, 1977  
 A>Title: Disulfide bonds in human serum albumin.  
 A;Reference number: A90930  
 A;Contents: annotation; disulfide bonds  
 R;Jacobsen, C.  
 Biochem. J. 171, 453-459, 1978  
 A>Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding  
 A;Reference number: A90299; MUID:7818630; PMID:656055  
 A;Contents: annotation; bilirubin-binding site  
 R;Peters, T.; Reed, R.G.  
 in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20,  
 A>Title: Serum albumin: conformation and active sites.  
 A;Reference number: A94408  
 A;Contents: annotation; binding sites  
 R;Harper, M.E.; Dugaiczky, A.  
 Am. J. Hum. Genet. 35, 565-572, 1983  
 A>Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes  
 A;Reference number: A90028; MUID:83279982; PMID:6192711  
 A;Contents: annotation; gene position  
 R;Walker, J.E.  
 FEBS Lett. 66, 173-175, 1976  
 A>Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.  
 A;Reference number: A46755; MUID:76257808; PMID:955075  
 A;Contents: annotation  
 A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid  
 R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
 FEBS Lett. 298, 266-268, 1992  
 A>Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphat  
 A;Reference number: A56294; MUID:92183881; PMID:1544460  
 A;Contents: annotation  
 A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p  
 ase activity  
 C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized  
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak  
 C;Comment: A large number of variants of human serum albumin have been described.  
 C;Genetics:  
 A;Gene: GDB:ALB  
 A;Cross-references: GDB:118990; OMIM:103600  
 A;Map position: 4q11-4q13  
 C;Superfamily: serum albumin; serum albumin repeat homology  
 C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox  
 F;1-18/Domain: signal sequence #status predicted <SIG>  
 F;19-24/Domain: propeptide #status experimental <PRO>  
 F;25-609/Product: serum albumin #status experimental <MPT>  
 F;29-202/Domain: serum albumin repeat homology <SA1>  
 F;166-174/Product: kinetinsin #status experimental <KIP>  
 F;221-394/Domain: serum albumin repeat homology <SA2>  
 F;413-592/Domain: serum albumin repeat homology <SA3>  
 F;27/Binding site: copper (His) #status predicted  
 F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4  
 F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 82.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 8,5e-74; Indels 0; Gaps 0;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHKEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 60  
 DB 25 DAHKEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPVB 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPVB 144

QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAACLLP 180  
 DB 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAACLLP 204  
 QY 181 KLDELROEGKASSAKORLKASLOKFGGERAFKAWARLSORFFPKABEAFVSKLVTDLT 240  
 DB 205 KLDELROEGKASSAKORLKASLOKFGGERAFKAWARLSORFFPKABEAFVSKLVTDLT 264  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPFLLEKSHCIAEVENDMPA 300  
 DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPFLLEKSHCIAEVENDMPA 324  
 QY 301 DLPSLAADFVESKOVCKNYABAKOVFLGMFLYFYARRHPDYSVLLILAKTYETTLBKC 360  
 DB 325 DLPSLAADFVESKOVCKNYABAKOVFLGMFLYFYARRHPDYSVLLILAKTYETTLBKC 384  
 QY 361 CAADAPHECYAKVDFEFPKPLVEEQNLKQNCLEFQOLGEYKFNQALVLRVTKKVPQVST 420  
 DB 385 CAADAPHECYAKVDFEFPKPLVEEQNLKQNCLEFQOLGEYKFNQALVLRVTKKVPQVST 444  
 QY 421 PTLVEVSRLKGVSGKCCCKHPKAKRMPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEVSRLKGVSGKCCCKHPKAKRMPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCPSALEVDETVYKPEFNAETFTFHADICTLSEKERQIKQOTALVELVKKHPKAT 540  
 DB 505 LVNRRPCPSALEVDETVYKPEFNAETFTFHADICTLSEKERQIKQOTALVELVKKHPKAT 564  
 QY 541 KEQLKAVNMDFAAFVEKCCCKADDDKTCFAEKGKLVAAASQAALGL 585  
 DB 565 KEQLKAVNMDFAAFVEKCCCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 2  
 A47391  
 serum albumin precursor - rhesus macaque  
 C;Species: Macaca mulatta (rhesus macaque)  
 C;Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C;Accession: A47391  
 R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
 A>Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilin  
 A;Reference number: A47391; MUID:93211971; PMID:8460152  
 A;Contents: B/B homozygote  
 A;Accession: A47391  
 A;Status: preliminary  
 A;Molecule type: mRNA; protein  
 A;Residues: 1-600 <WAT>  
 A;Cross-references: UNIPROT:Q28522; GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342293  
 A;Experimental source: liver  
 A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:128281)  
 C;Superfamily: serum albumin; serum albumin repeat homology  
 F;21-194/Domain: serum albumin repeat homology <SA1>  
 F;213-386/Domain: serum albumin repeat homology <SA2>  
 F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 77.7%; Score 2942; DB 2; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 1.4e-69;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 DAHKEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 60  
 DB 17 DTHKSEVAHFRKDLGEENFKGLVLAFAVSYLOQCPEFHVKLVNEVTEFAKTCVADESAAE 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPVB 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPVB 136  
 QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAACLLP 180  
 DB 137 DVMCTAFHDNEETFLKKYLIEIARRHPYFYFAPELLFFAKRYKAAFTCCQADKAACLLP 196

181 KLDLRDEGKASSAKORLKCASLQKFGERAFAKAWARLSQKFPKAEFAVSKLVTDLTK 240  
197 KLDLRDEGKASSAKORLKCASLQKFGDRAFAKAWARLSQKFPKAEFAVSKLVTDLTK 256  
241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 300  
257 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 316  
301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360  
317 DLPSLAADYVESKDVCKNYAEAKDVLGMLFYLYEYARRHPDYSVMLLRLAKAYEATLEKC 376  
361 CAADPHCEYAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 420  
377 CAADPHCEYAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 436  
421 PTLVEVSRNLGKVGSKCKPHEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
437 PTLVEVSRNLGKVGSKCKPHEAKRMPCAEDYLSVNLNQLCVLHEKTPVSEKVTCKCTES 496  
481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
497 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 556  
541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAAL 583  
557 KEQLKGVNDNFAPFVEKCKKADDKETCFABEGKLVAAASQAAL 599

RESULT 3  
S57632  
serum albumin precursor - cat  
C/Species: Felis silvestris catus (domestic cat)  
C/Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: JC4660; S57632  
R:Hilger, C.; Grigioni, F.; Hentges, F.  
Gene 165, 295-296, 1996  
A/Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A/Reference number: JC4660; MUID:96194824; PMID:8647469  
A/Accession: JC4660  
A/Molecule type: mRNA  
A/Residues: 1-608 <H12>  
A/Cross-references: UNIPROT:P49064; EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g8864  
A/Experimental source: liver  
C/Comment: This protein is the major protein component in plasma. It functions as a mult  
ein has 35 conserved cysteine residues.  
C/Superfamily: serum albumin, serum albumin repeat homology  
C/Keywords: liver; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRP>  
F:25-608/Product: serum albumin #status predicted <MAT>  
F:29-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 69.2%; Score 2620; DB 2; Length 608;  
Best Local Similarity 82.0%; Pred. No. 4.8e-61;  
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDILGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60  
Db 25 EAHQSEIAHRFNDLGEHFRLGLVLFVAFSQYLQCCPFEDHVKLVNEVTEFAKTCVADSGAA 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLRVPEV 120  
Db 85 NCKSLHELLGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMPGQGLVTP 144

Qy 121 DVMCTAFHDNEETFLKKVLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 145 DAMCTAFHENEQRFLGKLYEIARRHPYFYAPPELLFYAEEYKGVFTCECEAADKAACLLP 204

Qy 181 KLDLRDEGKASSAKORLKCASLQKFGERAFAKAWARLSQKFPKAEFAVSKLVTDLTK 240

205 KVDALREKVLASSAKERLKCASLQKFGERAFAKAWARLSQKFPKAEFAISKLVTDLAK 264  
241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 300  
265 IHKECCCHGDLLECCADRADLAKYICENQDSISTKLKECCCGPVLKSHCISEVERDELPA 324  
301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYLYEYARRHPDYSVVLLRLAKTYETTTLEKC 360  
325 DLPSLAADYVESKDVCKNYAEAKDVLGMLFYLYEYARRHPDYSVLLRLAKAYEATLEKC 384  
361 CAADPHCEYAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 420  
385 CATDDPPACVAHYDFEKPFLVEEPQNLIKONCELFQOLGEYKFNQALLVRYTKVPQVST 444  
421 PTLVEVSRNLGKVGSKCKPHEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
445 PTLVEVSRNLGKVGSKCKPHEAKRMPCAEDYLSVNLNQLCVLHEKTPVSEKVTCKCTES 504  
481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540  
505 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 564  
541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAAL 583  
565 KEQLKVMGDFGSDVKCCAAEDKEACFAEGPKLVAAQAAL 607

RESULT 4  
ABHOS  
serum albumin precursor - horse  
C/Species: Equus caballus (domestic horse)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C/Accession: S34053  
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.  
Eur. J. Biochem. 215, 205-212, 1993  
A/Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm  
A/Reference number: S34053; MUID:93345495; PMID:8344282  
A/Accession: S34053  
A/Molecule type: mRNA  
A/Residues: 1-607 <HOA>  
A/Cross-references: UNIPROT:P35747; GB:X74045; NID:g399671; PIDN:CAA52194.1; PID:g399672  
A/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, r  
teroid hormones (weak bonds with these hormones promote their transfer across the membr  
C/Superfamily: serum albumin, serum albumin repeat homology  
C/Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status predicted <MAT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,431  
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 65.4%; Score 2475.9; DB 1; Length 607;  
Best Local Similarity 76.3%; Pred. No. 3e-57;  
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

Qy 1 DAHKSEVAHRFKDILGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60  
Db 25 DTHKSEIAHRFNDLGEHFRLGLVLFVAFSQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 84

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLRVPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPKL-RPEP 143

Qy 121 DVMCTAFHDNEETFLKKVLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180  
Db 144 DAQCAAFQEDPDKFLGKLYEYARRHPYFYAPPELLFFAEYKADFTCECCPADDKAACLLP 203

Qy 181 KLDLRDEGKASSAKORLKCASLQKFGERAFAKAWARLSQKFPKAEFAVSKLVTDLTK 240

Db	204	KLDALKERILLSSAKERLKCSSFQNGFRVAVKAMSVARLSQKFPKADFAVSKIVTDLT	263
Qy	241	VHTECHGDLLECADRADLAKYICENQDSSISKLKCECEKPLEKSHCIAEVENDEMPA	300
Db	264	VHKECHGDLLECADRADLAKYICHQDSISGLKACCDKFLQKSHCIAEVKEDDLP	323
Qy	301	DLPSLAADFVSKGVCKVNAEKVFLGMLFYEARHPDYSVLLRLAKTYETTLK	360
Db	324	DLPALAADFADKCKYKAKOAVFLGTLFYESSRRHPDYSVLLRLAKTYETTLK	383
Qy	361	CAADPHCYAKVDFEPLVEEPQNLKONCELPQGLGEYKFNALLVRYTKVPQVST	420
Db	384	CAEADPPACRYTFVDFQPLVEEPKSLVKKNCDFEEVGEYDFQNALIVRYTKAPQVST	443
Qy	421	PTLVSVSNLGVSKCKKHPAKMPCAEADYLSVLNQLCVLHKTVPVSRVTKCTES	480
Db	444	PTLVIGRTLGVKGRCCCKLPESERLPCSENHALALNRLCVLHKTVPSEKITKCTDS	503
Qy	481	LVNRRPCFSALBVDYTVVPKFNAPETFTFHADICTLSKEROIKKQTLALVELVKKPKAT	540
Db	504	LAERRPCFSALBVDGYVPKFKPABTFTFHADICTLPDEKQIKKQSALAEVLVKKPKAT	563
Qy	541	KEOLKAVMDDFRAVFEKCKKADDKETCFRAEKGKLVASQAL	583
Db	564	KEOLKTVLGNFSAFVAKCCGREDKEACFAEAGPKLVASSQAL	606
RESULT 5			
ABBO5			
serum albumin precursor [validated] - bovine			
N:Alternate names: 67K protein; preproalbumin			
C:Species: Bos primigenius taurus (cattle)			
C:Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 09-Jul-2004			
C:Accession: A38885, A36401, A91258, B60808, S10780, D45800, A26693, A90309, A91458, A94			
R:Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.			
A:Submitted to the EMBL Data Library, August 1991			
A:Description: Bovine serum albumin: cDNA sequence and expression.			
A:Reference number: A38885			
A:Accession: A38885			
A:Molecule type: mRNA			
A:Residues: 1-607 <HOL>			
A:Cross-references: UNIPROT:P04277; EMBL:W73215			
R:Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.			
Biochem. Biophys. Res. Commun. 173, 639-646, 1990			
A:Title: Rapid confirmation and revision of the primary structure of bovine serum albumi			
A:Reference number: A36401; MUID:91083649; PMID:2260975			
A:Accession: A36401			
A:Molecule type: protein			
A:Residues: 25-41, 'H', '43-189, 'E', '191-213, 'T', '215-323, 'D', '325-393, 'TS', '396-607 <HIR>			
R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.			
Eur. J. Biochem. 98, 477-485, 1979			
A:Title: Biosynthesis of bovine plasma proteins in a cell-free system.			
A:Reference number: A91258; MUID:80024278; PMID:488109			
A:Accession: A91258			
A:Molecule type: protein			
A:Residues: 1-32 <MAG>			
R:Hsieh, J.C.; Lin, F.P.; Tam, M.F.			
Anal. Biochem. 170, 1-8, 1988			
A:Title: Electrophoresis onto glass-fiber filter from an analytical isoelectrofocusing			
A:Reference number: A60808; MUID:88267456; PMID:3389500			
A:Accession: B60808			
A:Molecule type: protein			
A:Residues: 25-41 <HSI>			
R:Strawich, E.; Glincher, M.J.			
Eur. J. Biochem. 191, 47-56, 1990			
A:Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is albu			
A:Reference number: S10780; MUID:90336641; PMID:2379503			
A:Accession: S10780			
A:Molecule type: protein			
A:Residues: 25-41, 'H', '43-57, 59-64 <STR>			
R:Carraway, R.E.; Cochran, W.; Mitra, S.P.			
J. Immunol. 143, 1680-1684, 1989			
A:Title: Structures of histamine-releasing peptides formed by the action of acid protease			

A:Reference number: A45800; MUID:89341406; PMID:2474609			
A:Accession: D45800			
A:Molecule type: protein			
A:Residues: 163-172 <CAR>			
R:Carraway, R.E.; Mitra, S.P.; Cochran, W.E.			
J. Biol. Chem. 262, 5968-5973, 1987			
A:Title: Structure of a biologically active neurotensin-related peptide obtained from pe			
A:Reference number: A26693; MUID:87194805; PMID:2437111			
A:Accession: A26693			
A:Molecule type: protein			
A:Residues: 165-172, 'L', 'CA2>			
R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.			
Biochem. J. 191, 867-868, 1980			
A:Title: Sequence of residues 400-403 of bovine serum albumin.			
A:Reference number: A90309; MUID:82023364; PMID:7283978			
A:Accession: A90309			
A:Molecule type: protein			
A:Residues: 402-433 <REE>			
R:Brown, J.R.			
Fed. Proc. 34, 591, 1975			
A:Title: Structure of bovine serum albumin.			
A:Reference number: A91458			
A:Accession: A91458			
A:Molecule type: protein			
A:Residues: 25-41, 'H', '43-117, 'EQ', '120-179, '181-189, 'E', '191-194, 'A', '196-213, 'T', '215-288, 'B'			
R:Brown, J.R.			
submitted to the Atlas, April 1975			
A:Reference number: A94551			
A:Accession: A94551			
A:Molecule type: protein			
A:Residues: 190-195 <BR2>			
R:Brown, J.R.			
Fed. Proc. 33, 1389, 1974			
A:Reference number: A91457			
A:Contents: annotation: disulfide bonds			
R:Werlen, R.C.; Offord, R.E.; Rose, K.			
Biochem. J. 302, 907-911, 1994			
A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC			
A:Reference number: S55232; MUID:95031935; PMID:7945219			
A:Accession: S55232			
A:Status: preliminary			
A:Molecule type: protein			
A:Residues: 529-536; 569-572 <WER>			
C:Superfamily: serum albumin; serum albumin repeat homology			
C:Keywords: carrier protein; copper binding; duplication; plasma			
F:1-18/Domain: signal sequence #status experimental <SIG>			
F:19-24/Domain: propeptide #status experimental <PRO>			
F:25-607/Product: serum albumin #status experimental <MPT>			
F:229-201/Domain: serum albumin repeat homology <SA1>			
F:220-393/Domain: serum albumin repeat homology <SA2>			
F:412-591/Domain: serum albumin repeat homology <SA3>			
F:27/Binding site: copper (His) #status predicted			
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,			
Query Match 64.6%; Score 2446.9; DB 1; Length 607;			
Best Local Similarity 75.6%; Pred. No. 1.8e-56;			
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;			
Qy	1	DAKSEVAHRFKDLGEEFKALVLITAFQYLQCCPFEDHVKLVNVEVTEPAKTCVADESAB	60
Db	25	DTKSEIAHRFKDLGEEFKGLVLITAFQYLQCCPFEDHVKLVNVEVTEPAKTCVADESHA	84
Qy	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPILRLVRPV	120
Db	85	GCEKSLHTLFGDELCKVASLRETYGDMADCCQEPERNECFLSHKDSDPLPKL-KPDP	143
Qy	121	DVMCTAFHDNETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAATECCQADKACLLP	180
Db	144	NTLCDKFADEKKFKYLYEIARHPYFYAPPELLFYANKYNGVFDCCQADKACLLP	203
Qy	181	KLDELDEGKASSAKORLCKASLOKFGFAPKAWAVARLSORFPKAEFAEVSKLVTDLTK	240
Db	204	KIETWREKVLASSARQORLCASIQKFGKALKAWVARLSQKFAEFVEVTKLVTDLTK	263

```
Qy 241 VHTCCGDLLECCADRADLAKYICENODSISKLKCCCKPILKESHCHIAEVENDEMPA 300
Db 264 VHKECCGDLLECCADRADLAKYICDNQDTLSSKLKCCCDPILKESHCHIAEVDKAVPE 323
Qy 301 DLPISLAADFVESKDVCKNYAEAKDVFGLMFYIYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 324 NLPLTADPAEDKDVCKNYQEAQAFGLSGFLYIYSRRHPYAVSVLLRLAKEYEATLEEC 383
Qy 361 CAADPHCEYAKVDFEFKPLVERPQNLKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420
Db 384 CAKDDPRACYSTVDFDKHLVDEPQNLKQNCDFEKLGEYGFQNALIVRYTRKVPQVST 443
Qy 421 PTLVEVSRNLGVSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 444 PTLVEVSRSLGVKTRCTKPESESRMPCDTEYLSLILNRLCVLHEKTPVSEKVTKCCTES 503
Qy 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKHKPAT 540
Db 504 LVNRRPCFSALTPTDETVVPKAFDEKLFTFHADICTLPTDEKQIKKQATLVELLKHKPKAT 563
Qy 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEKGKLVAAASQAAL 583
Db 564 EEQLKTMENFVAFVDKCAADDEKACFAVEGPKLVVSTQATL 606

RESULT 6
ABSHS
serum albumin precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S06936
R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A:Reference number: S06936; MUID:9009888; PMID:2602160
A:Accession: S06936
A:Molecule type: mRNA
A:Residues: 1-607 <BRO>
A:Cross-references: UNIPROT:P14639; EMBL:X17055; NID:gl186; PIDN:CAA34903.1; PID:gl1387
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-531/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 64.3%; Score 2432.9; DB 1; Length 607;
Best Local Similarity 75.0%; Pred. No. 4.1e-56;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

Qy 1 DAHKSEVAHRFKDGLGEENFKALVIAFAQYLQCPFFEDHVKLVNVEVTEFAKTCVADSESAE 60
Db 25 DTHKSEIAHRNFDLGEENFQGLVLIASFQYLQCPFFDEHVKLVKELTEFAKTCVADSESHA 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPFLRLVRPEV 120
Db 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCQEPERNECFNLNKHKQDSPDLKPKL-KPEP 143
Qy 121 DVMCTAFHDNETFLKYLVEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 144 DTLCAEFKADKKFWGKGLYIEVARHPYFYAPPELLYANKNGVFOBCCQEDKGACLLP 203
Qy 181 KLDELDRGKASSAKQRLKCSLQKFGGERFAKAWAVARLSQRFPAEFAVSKLVTDLTK 240
Db 204 KIDAMREKVLASSARQRLRCASIQKFGGERALKAWSVARLSQKFPKADFTDVTIVTDLTK 263
```

```
Qy 241 VHTCCGDLLECCADRADLAKYICENODSISKLKCCCKPILKESHCHIAEVENDEMPA 300
Db 264 VHKECCGDLLECCADRADLAKYICDHQDLSLSKLKCCCDPILKESHCHIAEVDKAVPE 323
Qy 301 DLPISLAADFVESKDVCKNYAEAKDVFGLMFYIYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 324 NLPLTADPAEDKDVCKNYQEAQAFGLSGFLYIYSRRHPYAVSVLLRLAKEYEATLEDC 383
Qy 361 CAADPHCEYAKVDFEFKPLVERPQNLKQNCLEFQOLGEYKFQNALLVRYTKVPQVST 420
Db 384 CAKDDPRACYSTVDFDKHLVDEPQNLKQNCLEFQOLGEYKFQNALLVRYTRKAPQVST 443
Qy 421 PTLVEVSRNLGVSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 444 PTLVEVSRSLGVKTRCTKPESESRMPCDTEYLSLILNRLCVLHEKTPVSEKVTKCCTES 503
Qy 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKKHKPAT 540
Db 504 LVNRRPCFSALTPTDETVVPKAFDEKLFTFHADICTLPTDEKQIKKQATLVELLKHKPKAT 563
Qy 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEKGKLVAAASQAAL 583
Db 564 DEQLKTMENFVAFVDKCAADDEKGCFCVLEGPKLVAASQAAL 606

RESULT 7
ABRTS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 09-Jul-2004
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:8122372; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SAR>
A:Cross-references: UNIPROT:P02770; GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; P
R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analyses
A:Reference number: A92211; MUID:77249657; PMID:893447
A:Note: cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
A:Residues: 25-222 <IS1>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino ac
A:Reference number: A91940; MUID:76260153; PMID:956149
A:Accession: A91940
A:Molecule type: protein
A:Residues: 223-288;572-608 <IS2>
A:Note: 262-Leu was also found
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; MUID:79001617; PMID:802665
A:Contents: annotation; copper binding
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid protease
A:Reference number: A45800; MUID:89341406; PMID:2474609
```

A:Accession: C45800  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 166-173 <AR>  
R:Heard, J.

Mol. Cell. Biol. 7, 2425-2434, 1987  
A>Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved  
A:Reference number: 157621; MUID:87286876; PMID:3475566

A:Accession: I57621

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:CROSS-references: GB:M16825; NID:g202828; PIDN:AAA0712.1; PID:g554412

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-18/Domain: signal sequence #status experimental <SIG>

F:19-24/Domain: propeptide #status experimental <PRO>

F:25-608/Product: serum albumin #status experimental <MAT>

F:29-202/Domain: serum albumin repeat homology <SA1>

F:221-394/Domain: serum albumin repeat homology <SA2>

F:413-592/Domain: serum albumin repeat homology <SA3>

F:27/Binding site: copper (His) #status experimental

F:77-86, 99-115, 114-125, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 340-385, 384-393, 4

Query Match 64.1%; Score 2426; DB 1; Length 608;

Best Local Similarity 73.4%; Pred. No. 6.3e-56;

Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 60

DB 25 EAHKSEIAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNDPNLPLRVPRPV 120

DB 85 NCDKSLHTLFGDKLCAIPSLREHYGDLADCCBPENECFLQHKDNDPNLPLRVPRPV 144

QY 121 DVMCTAFHNDNEFTFLKYLIEIARRHPYFYAPPELLYFAKRYKAAFTCCQAADKAACLTP 180

DB 145 EAMCTSFQENPTSFGLHYLHEVARHPYFYAPPELLYFAKRYKAAFTCCQAADKAACLTP 204

QY 181 KLDELDRDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

DB 205 KLDVAKKEKALVAARVMKCSSQRPGERAFKAWAVARMSQRPFAEVSCLVTDLT 264

QY 241 VHTCECHGDLLECCADRADLAKYICENODSISGLKECCPEKPLEKSHCHIAEVDNDMPA 300

DB 265 INKECHGDLLECCADRADLAKYICENODSISGLKECCPEKPLEKSHCHIAEVDNDMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEARHPDYSVLLRLRAKTYETTLK 360

DB 325 DLPSIAADFVEDKEVCKNYAEAKDVFLGMFLYEARHPDYSVLLRLRAKTYETTLK 384

QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLIKQNCLEFQGEYKFONALLVRYTKVPQVST 420

DB 385 CAEGDPACVGTVLAEFQPLVEEPQNLIKQNCLEFQGEYKFONALLVRYTKVPQVST 444

QY 421 PTLVEVRNLGKVGSKCCCHPKAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480

DB 445 PTLVEARNLGRVGTAKCCTLPKQRLPCVEDYLSATLNLRLCVLHEKTPVSEKVKCCSGS 504

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLALVELVKKPKAT 540

DB 505 LVRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLALVELVKKPKAT 564

QY 541 KQOLKAVMDDFAFVBEKCKADDKETCFABEGKLVAAQAL 583

DB 565 EDQLTKVMGDFQFVDKCKCAADKDCNCFATEGNLVARSKAL 607

RESULT 8

ABPGS

serum albumin precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C:Accession: S01382; A61006

R:Weinstock, J.; Baldwin, G.S.

Nucleic Acids Res. 16, 9045, 1988

A>Title: Nucleotide sequence of porcine liver albumin.

A:Reference number: S01382; MUID:89016582; PMID:3174440

A:Accession: S01382

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-605 <WEI>

A:CROSS-references: UNIPROT:P08835; EMBL:X12422; NID:g1875; PIDN:CAA30970.1; PID:g8333798

R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.

J. Bone Miner. Res. 4, 235-241, 1989

A>Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral

A:Reference number: A61006; MUID:89269769; PMID:2728927

A:Accession: A61006

A:Molecule type: protein

A:Residues: 23-51, 'X', 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>

A:Experimental source: dental enamel

A>Note: albumin and other serum proteins are also found in bone

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,

teroid hormones (weak bonds with these hormones promote their transfer across the membra

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>

F:17-22/Domain: propeptide #status predicted <PRO>

F:23-605/Product: serum albumin #status predicted <MAT>

F:27-199/Domain: serum albumin repeat homology <SA1>

F:218-391/Domain: serum albumin repeat homology <SA2>

F:410-589/Domain: serum albumin repeat homology <SA3>

F:75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390,

F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 63.9%; Score 2416.8; DB 1; Length 605;

Best Local Similarity 71.8%; Pred. No. 1.1e-55;

Matches 441; Conservative 70; Mismatches 71; Indels 32; Gaps 2;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 60

DB 23 DTYKSEIAHRFKDLGSENFKALVLIAPAYLOQCPEFHVKLVNEVTEPAKTCVADESAB 82

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNDPNLPLRVPRPV 120

DB 83 NCDKSLHTLFGDKLCAIPSLREHYGDLADCCBPENECFLQHKDNDPNLPLRVPRPV 141

QY 121 DVMCTAFHNDNEFTFLKYLIEIARRHPYFYAPPELLYFAKRYKAAFTCCQAADKAACLTP 180

DB 142 VALCADFOEDGQKFWGKYLIEIARRHPYFYAPPELLYFAKRYKAAFTCCQAADKAACLTP 201

QY 181 KLDELDRDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

DB 202 KIEHLREKVLTSAAQRLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 261

QY 241 VHTCECHGDLLECCADRADLAKYICENODSISGLKECCPEKPLEKSHCHIAEVDNDMPA 300

DB 262 VHTCECHGDLLECCADRADLAKYICENODSISGLKECCPEKPLEKSHCHIAEVDNDMPA 321

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEARHPDYSVLLRLRAKTYETTLK 360

DB 322 DLNPLEHDFVEDKEVCKNYAEAKDVFLGMFLYEARHPDYSVLLRLRAKTYETTLK 381

QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLIKQNCLEFQGEYKFONALLVRYTKVPQVST 420

DB 382 CAKEDPPACVATVDFKQPLVDEPKNLIKQNCLEFQGEYKFONALLVRYTKVPQVST 441

QY 421 PTLVEVRNLGKVGSKCCCHPKAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480

DB 442 PTLVEVARNLGLVGRCKRPEERLSCAEDYLSVLNQLCVLHEKTPVSEKVKCCCTES 501

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLALVELVKKPKAT 540

DB 502 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLALVELVKKPKAT 561

```
Qy 541 KEQLKAYMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGLMSPLRLEVPCHALPQ 600
Db 562 EEQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKF-----596
Qy 601 GLSPGQVIVRGLV 614
Db 597 -----VIEIRGIL 604

RESULT 9
JC5838
serum albumin - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5838
R;Yoshida, K.; Sero-Oshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the
A;Reference number: JC5838; MUID:98116663; PMID:9455485
A;Accession: JC5838
A;Molecule type: mRNA
A;Residues: 1-609 <YOS>
A;Cross-references: UNIPROT:O35090; DBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2
A;Experimental source: liver
C;Superfamily: serum albumin; serum albumin repeat homology
F;222-395/Domain: serum albumin repeat homology <SA2>

Query Match 63.1%; Score 2387; DB 2; Length 609;
Best Local Similarity 73.9%; Pred. No. 6.8e-55;
Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

Qy 2 AHKSEVARRFKDLGEENFKALVIAFAQYLQOCPEFDHVKLVNVEVTRFAKTCVADSAEN 61
Db 27 AHKSEIAHRYKDLGEKFKGLVLVTFQYLQKCYEEHVKLVREVTDFASNCAKDESAEN 86
Qy 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVLRPEVD 121
Db 87 CDKSLHTLFGDKLCSLNFGEKYAEMADCCAKQEPERNECFLOHKDDNPQLPPFKRAEPD 146
Qy 122 VMCTAFHDNETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLPK 181
Db 147 AMCTAFQENAEAFMGHYLHEVARRHPYFYGPPELLYLDKVTAVLTGECQAADKGCALTPK 206
Qy 182 LDELDEGKASSAKORLKCSAQKQGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTKV 241
Db 207 LDALKEKALVSARQRLKCSSMKKFGERAFKAWAVARMSQTFPNADFAEITKGLATDLTKV 266
Qy 242 HTECHGDLLECADDRLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPAD 301
Db 267 TQECHGDLLECADDRLAKYKMCENQASISSKLQACCDKEMLQKSQLAEVEHDDMPAD 326
Qy 302 LPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLRLAKTYETTLBKCC 361
Db 327 LPALTADFVEDKDVCKNYAEAKDVLGTFLYEYSRRHPDYSVSLLRLAKKYEATLEKCC 386
Qy 362 AAADPHCYAKVDFEFLVEEPQNLKQNCLEFQIGYKFNQALLVRYTKKVPQVSTP 421
Db 387 AEADPHACYGHVDFEFLVEEPQNLVKSNCLEKYGKFNQALLVRYTKKAPQVSTP 446
Qy 422 TLVEVSRLNGKVGSKCKHPKAMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCYESL 481
Db 447 TLVEAARSLGRVTHCCALPEKKRLPCVEDYLSAILNRCVLLHKEKTPVSEQVTKCCSGSL 506
Qy 482 VNRPCFSALVEVDYTPKFNFAETFTFHADICTLSEKERQIKKQATLVELVHKHPKATK 541
Db 507 VERRPCFSALPVDYTPKFKFAETFTFHANICTLPEKEKQMEKQATALAEVLVHKHPQATE 566
Qy 542 BQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAAL 583
Db 567 BQLKAVMGDFAEFLEKCKQEDKEACFSTEGPKLVAESQKAL 608

RESULT 10
```

```
A05139
serum albumin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: A05139; I48638
R;Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A;Reference number: A93055; MUID:88216123; PMID:2452956
A;Accession: A05139
A;Molecule type: mRNA
A;Residues: 1-418 <MIN>
A;Cross-references: UNIPROT:P07724; GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R;Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A;Reference number: I48638; MUID:90289606; PMID:1971802
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 379-453 <BOC>
A;Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology <SA2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 49.2%; Score 1861; DB 2; Length 453;
Best Local Similarity 72.2%; Pred. No. 2e-41;
Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;

Qy 75 CTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVPEVDVMCTAFHDNETF 134
Db 1 CALPNLRKENTGELADCCCTKQEPERNECFLOHKDDNPSPLPFPERPEAMCTSKENFTTF 60
Qy 135 LKLYLYEIAARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLPKLDELDEGKASSA 194
Db 61 MGHYLHEVARRHPYFYAPPELLYLYAEQYNEILTQCCAEADKESCLTPKLDGVKEKALVSSV 120
Qy 195 KQRLKASLQKQGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTKVHTCCCHGDLLECA 254
Db 121 RQRMKSSMQKQGERAFKAWAVARLSQTFPNADFAEITKGLATDLTKVNCCKCHGDLLECA 180
Qy 255 DDLADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPADLPSLAADFVESKD 314
Db 181 DDLAEALAKYICENQATISSKLTQCCDKPLKKAHCLSEVEHDTMPADLPAIADFVEDQE 240
Qy 315 VCKNYAEAKDVLGMFLYFYARRHPDYSVVLLRLAKTYETTLKCCAAADPHCYAKVF 374
Db 241 VCKNYAEAKDVLGTFLYEYSRRHPDYSVSLLRLAKKYEATLEKCCAEANPPACYGTVL 300
Qy 375 DEFKPLVEEPQNLKQNCLEFQIGYKFNQALLVRYTKKVPQVSTFTLVEVSRLNGKVG 434
Db 301 AEFQPLVEEPKLVKTNCDLYEKLGEYGFQNAIILVRYTKAPQVSTFTLVEAARNLGRVG 360
Qy 435 SKCKCHPEAKRMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCYESLVNRRPCFSALVEVD 494
Db 361 TKCCTLPEDQRLPCVEDYLSAILNRCVLLHKEKTPVSHVTKCCGSLVERRPPCFALSALTVD 420
Qy 495 ETVPKPFNAETFTFHADICTLSEKERQIKKQT 527
Db 421 ETVPKPFNAETFTFHSDICTLPEKEQIKKQT 453

RESULT 11
ABCS8
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S15571; A05078; A13451
R;Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
```









QY 352 TYETTLKCCAAADPHCYAKVDFEKLVPBPQNLKQNCLEPQLGEYKFNALLVRY 411  
 Db 376 GYQELLEKCFQTNPLECQDKEEELQKYIOESQALAKRSCLGFKLGEYLYQNAFLVAY 435  
 QY 412 TKKYPQVSTPTLVEVSRNLGVSKCKHPKAPKMPCAEDYLVSVNLQCLVHKEKTPVSD 471  
 Db 436 TKKAPQLTSSSELMAITRMAATAATCCQSEDKLLACGEGAADIIGHLCIRHEWTPVNP 495  
 QY 472 RVTKCTESLVNRRPCFSALEVDETYVPKFNAAETFTPHADICTLSEKERQIKKQTALVE 531  
 Db 496 GVGQCTSSYANRRPCFSLVDETYVPAPFSDDKFIFHKOLCOAQGVALQTMKQEFLLN 555  
 QY 532 LVKHKPKATKXQLKAVMDDFAFVKECKCKADDDKTCFAEKGKLVAAQAALGL 585  
 Db 556 LVKQKPOITEQLEAVIADFSLLEKCCQGOEVCFAEKGKLSKTRALGV 609

RESULT 14  
 FPGO  
 alpha-fetoprotein precursor - gorilla  
 C:Species: Gorilla gorilla (gorilla)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A37970  
 R:Ryan, S.C.; Zielinski, R.; Dugaiczky, A.  
 Genomics 9, 60-72, 1991  
 A:Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates.  
 A:Reference number: A37970; MUID:91169517; PMID:1706310  
 A:Accession: A37970  
 A:Molecule type: DNA  
 A:Residues: 1-609 <RYA>  
 A:Cross-references: UNIPROT:P28050; GB:M38272; NID:9817963; PIDN:AAA73520.1; PID:9177041  
 C:Genetics:  
 A:Map position: 4q11-12  
 A:Introns: 29/1; 46/2; 205/3; 161/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-609/Product: alpha-fetoprotein #status predicted <MAT>  
 F:29-202/Domain: serum albumin repeat homology <SA1>  
 F:221-394/Domain: serum albumin repeat homology <SA2>  
 F:413-592/Domain: serum albumin repeat homology <SA3>  
 F:22/Binding site: copper (His) #status predicted  
 F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472  
 P:249/Binding site: bilirubin (Lys) #status predicted  
 F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.9%; Score 1245.7; DB 1; Length 609;  
 Best Local Similarity 39.4%; Pred. No. 9.1e-25;  
 Matches 234; Conservative 115; Mismatches 232; Indels 13; Gaps 3;

QY 4 KSEVAHR-----PKDLGEENFKALVLIAPQYLQCPFDHVKLVNEVTEFAKT 52  
 Db 17 ESRTLHRNRYGSIADSDYQTAELSLADATIFFAQFVQEAITYEVSXKWDALTAIBK 76  
 QY 53 CVADSEAKCNSLHFLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNP-N 111  
 Db 77 PTGDSQACLENQLPAFLLEELCHEILEKYG-LSDDCSQSEEGHNCFLAKHPTPAS 135  
 QY 112 LPLRVPEVDVMTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKYKAAFTCCQA 171  
 Db 136 IPLFQVPEPVTSCAEVEDRETFMNFYIEIARRHPFYAPTILLWAARYDKIIFSCCKA 195  
 QY 172 ADKAACLLPKLDELDEGKASSAKQLKASLQKGERAFKAWAVARLSORPKAEFAV 231  
 Db 196 ENAVECFQTKAATVTVKELRESLLNHACAVMKNFGTRTFQAITVTKLSQKFTKNFTBI 255  
 QY 232 SKLVTLTKVHTBCCHDLECCADDRADLAKYICENQDSSIKLKECKEPLLEKSHCIA 291  
 Db 256 QKLVLDAVHVEHCCRGDVLDCLODEKIMSYICSDQDTLSNKTIECCKLTLLERQCII 315  
 QY 292 EVENDEMPADLPSSAADPVESKDVCKNYAAKDVFLGMFLMEYARRHPYVSVLLRLAK 351  
 Db 316 HAENDEKPEGLSPNLRFLGDRDFNQFSSGKKNIFLASFVHEYSRRHPQLAVSVILRVAK 375

QY 352 TYETTLKCCAAADPHCYAKVDFEKLVPBPQNLKQNCLEPQLGEYKFNALLVRY 411  
 Db 376 GYQELLEKCFQTNPLECQDKEEELQKYIOESQALAKRSCLGFKLGEYLYQNAFLVAY 435  
 QY 412 TKKYPQVSTPTLVEVSRNLGVSKCKHPKAPKMPCAEDYLVSVNLQCLVHKEKTPVSD 471  
 Db 436 TKKAPQLTSSSELMAITRMAATAATCCQSEDKLLACGEGAADIIGHLCIRHEWTPVNP 495  
 QY 472 RVTKCTESLVNRRPCFSALEVDETYVPKFNAAETFTPHADICTLSEKERQIKKQTALVE 531  
 Db 496 GVGQCTSSYANRRPCFSLVDETYVPAPFSDDKFIFHKOLCOAQGVALQTMKQEFLLN 555  
 QY 532 LVKHKPKATKXQLKAVMDDFAFVKECKCKADDDKTCFAEKGKLVAAQAALGL 585  
 Db 556 LVKQKPOITEQLEAVIADFSLLEKCCQGOEVCFAEKGKLSKTRALGV 609

RESULT 15  
 ABXL72  
 74K albumin precursor - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: B41682; S02693; A05288  
 R:Moskaltis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.  
 Mol. Endocrinol. 3, 464-473, 1989  
 A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic a  
 e during development.  
 A:Reference number: A41682; MUID:89313788; PMID:2747653  
 A:Accession: B41682  
 A:Molecule type: mRNA  
 A:Residues: 3-607 <MOS>  
 A:Cross-references: UNIPROT:P14872; GB:M21442; NID:9213930; PIDN:AAA49637.1; PID:9213931  
 R:Schorp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U.  
 J. Mol. Biol. 199, 83-93, 1988  
 A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Del  
 A:Reference number: S02692; MUID:86172470; PMID:2451026  
 A:Accession: S02693  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-48 <SCH>  
 A:Cross-references: EMBL:226826  
 R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J  
 Eur. J. Biochem. 146, 489-496, 1985  
 A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilizati  
 A:Reference number: A05288; MUID:85126974; PMID:3971963  
 A:Accession: A05288  
 A:Molecule type: mRNA  
 A:Residues: 459-502, 'L', 504-557 <WOL>  
 A:Cross-references: GB:M28276  
 A>Note: the authors translated the codon TAT for residue 63 as Thr  
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,  
 mones (weak bonds with these hormones promote their transfer across the membranes), thy  
 C:Genetics:  
 A:Introns: 27/1  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-24/Domain: propeptide #status predicted <PRO>  
 F:25-607/Product: 74K serum albumin #status predicted <MAT>  
 F:32-201/Domain: serum albumin repeat homology <SA1>  
 F:220-393/Domain: serum albumin repeat homology <SA2>  
 F:412-591/Domain: serum albumin repeat homology <SA3>  
 F:30/Binding site: copper (His) #status predicted  
 F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392  
 F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.9%; Score 1206.6; DB 1; Length 607;  
 Best Local Similarity 39.3%; Pred. No. 9.7e-24;  
 Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

QY 3 HKSEVAHRPKDLGEENFKALVLIAPQYLQCPFDHVKLVNEVTEFAKTVADESAENC 62

```

Db      30  HHKHIADVYTALTERTFKGLTLAIVSQNLQKCSLEELSKLVWEINDFAKSCINDKTPE-C 88
Qy      63  DKSLSHTLFGDKLCTVATILRETYGEMADCCAKQOEPRNECFLOHKDDNPNIPLRVPRVEDV 122
Db      89  EKPVGTLFFDKLCADPAVGNYEWSKECCAKQDPERAQCFKAHRDHBT---SIKPEPFE 145
Qy     123  MCTAFHDSNEETFLKKLYEYARRHPYFVAPELLFFAKRYKAAFTPECCOAAADKAACLLPKL 182
Db     146  TKLLKKEHPDDLISAFIHEARNHPDLYPPAVALTATKQYHKLAHCCBEEDKEKCFSEKM 205
Qy     183  DELRDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFPKAEFAEVSCLVTLTKVH 242
Db     206  KQLMKQSHSTEDKQHFECWILDNFPEKVLKALNARVSHRYPKAEFLAHNFTVEVTHFI 265
Qy     243  TECCHGDLLECADRADLAKYICENQDSISSKLKECCKEPILLESKHCIAEVENDEPADL 302
Db     266  KDCCHDDMFECMTERBELTECHQKDELSEKLECCNIPLLERTYICVITLENDVDVPAEL 325
Qy     303  PSLSAADFVESKDVCKNYAEAKDVFGLGMFLYEAARRHPDYSVLLRLRAKTYETTLEKCCA 362
Db     326  SQPTETEDPHVCEKKAENNEVFLGRYLHAVSRKHQELSEQFLQSAKEYESLLNKCK 385
Qy     363  AADPHECYAKVDFEFKPLVEPQNLIKONCELFQQLGEYFQNALLVRYTKKVPQVSTPT 422
Db     386  TDNPPECYKQAGDRFMNEAKERFAYLKQNCDDILHEHGYYLFENELLIRYTKKMPQVSD 445
Qy     423  LVEVSRNLKGVKSCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLV 482
Db     446  LIGIAHQWADIGEHCCCAVPENQRPFCAEGLDTILIGKMCERQKKTFFNNHVAHCCDTSYS 505
Qy     483  NRRPCFSALEVDENYVVPKFAEFTFHADICTLSEKERQIKKOTALVELVGHKPKATKE 542
Db     506  GMRSCFTALGDEDDYVPPVTDTHFFDDKICTANDKEKHQIKQKFLVLKLVKSPKLEKN 565
Qy     543  QLKAVMDDFAAVFVKCKCKADDKETCFAEEGKLVAAASQ 580
Db     566  HIDECSAEFLKMWQCKCTADEHQCFDTEKPVLTIEHQ 603

RESULT 16
I46986
albumin - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999
C:Accession: I46986
R:Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner
J. Allergy Clin. Immunol. 93, 614-627, 1994
A:Title: Molecular characterization of dog albumin as a cross-reactive allergen.
A:Reference number: I46986; UID:94201492; PMID:7512102
A:Accession: I46986
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-265 <SPI>
A:Cross-references: GB:S72946; NID:G633937; PIDN:AA30434.1; PID:G633938
C:Superfamily: serum albumin; serum albumin repeat homology
F:7-180/Domain: serum albumin repeat homology <SA2>

Query Match      31.2%; Score 1181.9; DB 2; Length 265;
Best Local Similarity 82.6%; Pred. No. 3.1e-24;
Matches 218; Conservative 23; Mismatches 22; Indels 1; Gaps 1;

Qy     192  SSAQRLKASLOKQGERAFKAWAVARLSORFPKAEFAEVSCLVTLTKVTECHGDL 251
Db      2  SSAKERFKASLOKQGERAFKAWAVARLSORFPKADFAETSKVVTDLTKVHKECCHGDL 61
Qy     252  ECADRADLAKYICENQDSISSKLKECCKEPILLESKHCIAEVENDEPADLPSLSAADPVE 311
Db     62  ECADRADLAKYICENQDSISTKLKECCDKVPLEKSQCLAEVERDELPGDLPLSAADPVE 121
Qy     312  SKDYCKNYAEAKDVFGLGMFLYEAARRHPDYSVLLRLRAKTYETTLEKCCAAADPHECYA 371
Db     122  DKEVCKNYQAEAKDVFGLGTFLEYEYARRHPDYSVLLRLRAKTYETTLEKCCATDDPPTCYA 181
```

```

Qy     372  KVFDEFKPLVEEPQNLIKONCELFQQLGEYFQNALLVRYTKKVPQVSTPTL-VEVERNL 430
Db     182  KVLDEFKPLVDEPQNLVKTNCELFQQLGEYFQNALLVRYTKKAPQVSTPTLVVEVSRKL 241
Qy     431  GKVSGKCKCKHPEAKRMPCAEDYLS 454
Db     242  KGVGTCKCKKPESEMSCADDFLS 265

RESULT 17
ABXL68
68K serum albumin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A41682; S02692
R:Mockaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
e during development.
A:Reference number: A41682; UID:89313788; PMID:2747653
A:Accession: A41682
A:Molecule type: mRNA
A:Residues: 1-608 <MOS>
A:Cross-references: UNIPROT:P08759; GB:M18350
R:Schorpp, M.; Doebbeling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Dele
A:Reference number: S02692; UID:88172470; PMID:2451026
A:Accession: S02692
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:Z26825
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thy
C:Genetics:
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-608/Product: 68K serum albumin #status predicted <MAP>
F:32-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89,102-118,117-128,148-193,224-270,269-277,289-303,302-313,340-385,384-393,

Query Match      31.1%; Score 1178.3; DB 1; Length 608;
Best Local Similarity 38.3%; Pred. No. 5.5e-23;
Matches 222; Conservative 111; Mismatches 240; Indels 7; Gaps 2;

Qy      3  HKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPEPDHVKLVNVEVTEFAKTCVADESANC 62
Db     30  HHKHIADVYNLLTERTFKGLTLAIVSQNLQKCSLEELSKLVWEINDFAKSCINDKTPEC 89
Qy     63  DKSLSHTLFGDKLCTVATILRETYGEMADCCAKQOEPRNECFLOHK--DDNPNIPLRVPR 120
Db     90  EKPVGTLFFDKLCADPAVGNYEWSKECCSKQDPERAQCFKAHRVFEHP-----VRPK 144
Qy    121  DVMCTAFHDSNEETFLKKLYEYARRHPYFVAPELLFFAKRYKAAFTPECCOAAADKAACLLP 180
Db    145  EETCALFKHEHPDDLISAFIHEARNHPDLYPPAVALLTQYQKLVHEHCCBEEDKCKFAE 204
Qy    181  KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFPKAEFAEVSCLVTLTK 240
Db    205  KMKELMKHSHSIEDKQKFCWVNNYPERVVKALNARVSHRYPKDPFKLAHKFTEETH 264
Qy    241  VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCKEPILLESKHCIAEVENDEMPA 300
Db    265  FIKDCCHGDMFECMTERLEUSEHTECHQKDELSTKLKCCNIPLLERTYICVITLENDVDPA 324
Qy    301  DLPSLSAADFVESKDVCKNYAEAKDVFGLGMFLYEAARRHPDYSVLLRLRAKTYETTLEKC 360
```







N;Alternate names: alpha-albumin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C;Accession: A53195  
R;Belanger, L.; Roy, S.; Allard, D.  
J. Biol. Chem. 269, 5481-5484, 1994  
A;Title: New albumin gene 3' adjacent to the alpha-1-fetoprotein locus.  
A;Reference number: A53195; MUID:94164881; PMID:7509788  
A;Accession: A53195  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-608 <BEL>  
A;Cross-references: UNIPROT:P36953; GB:X76456; NID:g456358; PIDN:CAA53994.1; PID:g456358  
C;Genetics:  
A;Introns: 30/1; 47/2; 91/3; 162/2  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: glycoprotein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-608/Product: afamin #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-590/Domain: serum albumin repeat homology <SA3>  
  
Query Match 24.6%; Score 932; DB 2; Length 608;  
Best Local Similarity 33.7%; Pred. No. 1.7e-16;  
Matches 192; Conservative 104; Mismatches 264; Indels 10; Gaps 4;  
  
Qy 17 ENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCT 76  
Db 41 ENVAYLIIASAQVQEAESFEVEMLVKMDLYDKRCLADSTLPECKSIANDAIQDMLCD 100  
  
Qy 77 VATLRETYGEMADCAKQEPERNECFLOHKDDN----PNLPRVPEVDVWCTAFHNEE 132  
Db 101 MKGLPQKH-NFSHCRCQAQFORLCCFFYNKCANVGLFPFPPTL---DPEERCKQAYKNSE 156  
  
Qy 133 TFLKYLIEIARRHPFYFAPPELLFPAKRYKAAFTCCQAADKAACLPLKDLDEDEKAS 192  
Db 157 SFLNLYMEVARRPFPAPVLLNVAARFEAAATTCCEQQOKATYFODKAAPIYQYLKAL 216  
  
Qy 193 SAKQRLKCSAQKQGERAFKAWAVARLSORFPKAEVSKLVTDLTQVHTCECHGDLLE 252  
Db 217 SSYQRNVCGALLKFGPKTLNINIAVFSKFKPGIGFEDLTSLLEDVSSMYDGCCEGDVVQ 276  
  
Qy 253 CADRADLAKVICNQDISISKLCECKECPLEKSHCIAEVENDEMPADLPSLAADFVES 312  
Db 277 CIRSQSQVMHHICSKQDISISKIRACCEKLPERADCIINANKODRPEDLSLRTPKFTDS 336  
  
Qy 313 KDVCKNYAAEKDVLGMLFYVARRHPDYSVVLRLRLAKTYETTLKCCAAADPHECVAK 372  
Db 337 ENVQERDSQDKFAEFYDYSRRHSTLPELLRITKYKDLLEDCCNKNPLSCVYR 396  
  
Qy 373 VFDEFKPLVEFPQNLIKQCELFQGLGRYKFNALLVRYTKVPQVSTFTLVEVSRNLGK 432  
Db 397 AEDKFNETTESLAWVQCEQFQELGKDALQRHFLVKFKAAPQLPMEELVSLSKENVA 456  
  
Qy 433 VGSKCKRHPEAKRMPCABDYLSVLNLQCLVLEHKTTPVSDRVTKCTESLVNRRPCFSALE 492  
Db 457 ALATCCTLSL--EPACVDNLADVLGLCGMKNKRTINPTVHCCRADFAFRRPCFEHLK 514  
  
Qy 493 VDEITYVPKFNAYETFFHADICTLSEKEROIKQOTALVELVKHKPKATKEQLKAVMDQFA 552  
Db 515 ADITYALPSVALSALRADMCQPLKEDLQNRHRLVNLVKWMPETITDEERLCLFTKFT 574  
  
Qy 553 AFVEKCCKADDKETCFABEGKKLVAASQAA 582  
Db 575 AAGECGNIQKPEACFSPESSKTGDVSDQA 604

RESULT 23  
ABONS1  
serum albumin 1 precursor - Atlantic salmon  
C;Species: Salmo salar (Atlantic salmon)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C;Accession: A36238; S13079  
R;Byrnes, L.; Gannon, F.  
DNA Cell Biol. 9, 647-655, 1990  
A;Title: Atlantic salmon (Salmo salar) serum albumin: cDNA sequence, evolution, and tissue  
A;Reference number: A36238; MUID:91083837; PMID:2261082  
A;Accession: A36238  
A;Molecule type: mRNA  
A;Residues: 1-608 <BYR>  
A;Cross-references: UNIPROT:P21848; GB:X52397; NID:g64375; PIDN:CAA36643.1; PID:g64376  
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, iron  
mones (weak bonds with these hormones promote their transfer across the membranes), thyroxine  
C;Superfamily: serum albumin; serum albumin repeat homology  
C;Keywords: carrier protein; duplication; plasma  
F;1-14/Domain: signal sequence #status predicted <SIG>  
F;15-18/Domain: propeptide #status predicted <PRO>  
F;19-608/Product: serum albumin #status predicted <MAT>  
F;23-196/Domain: serum albumin repeat homology <SA1>  
F;215-390/Domain: serum albumin repeat homology <SA2>  
F;411-591/Domain: serum albumin repeat homology <SA3>  
F;26-72,71-80,93-108,107-118,142-187,186-195,218-264,263-271,283-299,298-309,336-381,380  
  
Query Match 20.2%; Score 765.3; DB 1; Length 608;  
Best Local Similarity 27.3%; Pred. No. 4.4e-12;  
Matches 167; Conservative 118; Mismatches 260; Indels 67; Gaps 12;  
  
Qy 4 KSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAENCD 63  
Db 22 QNCTITFTAKEDGFKSLILVGLAQNLPDSTLGLVPLIAEALAMGVKCCSDTPPEDCE 81  
  
Qy 64 KSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVLR-----P 118  
Db 82 RDVADLFSQAVCSSETLVEK-NDLKMCCEKTAERTHCFVDHK---AKIPRDLSLKAEPL 137  
  
Qy 119 EVDVMCTAFHDNEETFLKYLIEIARRHPFYFAPPELLFPAKRYKAAFTCCQAADKAACL 178  
Db 138 AAD-QCEDFKDKHKAQVGRFIFKFSKSNPMLPPHVLAIAGYGEVLVITTCGGEAAQTCF 196  
  
Qy 179 LPKLDELDRDGKAS-----SAKQRLKCSAQKQGERAFKAWAVARLSORFPKAEFAE 230  
Db 197 -----DTKCATFQHAVMKRVAELRSLCLIVHKYIGDRVYKAKLVQYSSQMPQASFOE 248  
  
Qy 231 VSKLVTDLTKVHTCCCHGDLLECCADRADLAKYICENQDISIS--SKLKECCEKPLEKSH 288  
Db 249 MGNVNDKIVATVAPCCSGDMVTCMKERKTLVDEVCADESVLSSRAAGLSACCKEDAVHRGS 308  
  
Qy 289 CIAEVENDEMP-----ADLPSLAADFVESKQVCKNYAAEKDVLGMLFYVARRH 338  
Db 309 CVEAMKDPKPDGLSELSEHYDIHADIAA-----VCQTFKTPDVMAGKLVYEISVRH 358  
  
Qy 339 PDYSVLLRLAKTYETTLKCCAAADPHECV-----AKVDEFKPLVEFPQNL 388  
Db 359 PESSQOVLTRFAKEAEQALQCCDMEDHAECCVKTALAGSDIDKKITDE-----TDYY 410  
  
Qy 389 KQCELFQGLGRYKFNALLVRYTKVPQVSTFTLVEVSRNLGKVGSKCKRHPEAK-RMP 447  
Db 411 KMCACAAAVSDDSFESKSMVYVYTRIMPQASFDQLHMVSEIVHDVLHACCKDEQGHFVLP 470  
  
Qy 448 CAEDYLSVVLNLQCLVLEHKTTPVSDRVTKCTESLVNRRPCFSALEVDEITYVPKFNASTFF 507  
Db 471 CAEKLDAIDATCDDYDPSPINPHIAHCCNQSYSMRRHCILAIQPDTEFTFPPELDASSF 530  
  
Qy 508 TFHADICTLSEKEROIKQOTALVELVKHKPKATKEQLKAVMDQFAAFVEKCKCKADDKETC 567  
Db 531 HMGPELCTKDSKDLLSGKLLYGVVRHKTITTEDHLTKITSTKYTHMKKCAAEQAAAC 590  
  
Qy 568 FAEEGKKLVAAS 579  
Db 591 FTEAPKLVSSES 602

RESULT 24  
ABONS2  
serum albumin 2 precursor - Atlantic salmon













QY 76 -----TVATRETYGEMADCCAKQEPERNECFLOHKDDN----- 109  
Db 272 DITLHNHSAALAAVVPTLTFLAGFGDC-----NLSDNEGQVMSAGWTEE 318  
QY 110 -----PNLPRLVRPEVDVYMCCTAFHDNEETFLKK 137  
Db 319 PVALIORMLFRTVLHMLSVDSMAEVMPENLRKNTETELRAALKI-----RTFLEK 369  
QY 138 YLYEIARHPYFAPPELLFFAKRYKA-----APTECCOQADKAACLLPKLDELDRDE 188  
Db 370 -----QDPD--FAPRQKTLQEVQDDFVFSKYCH-----RVLLLPELLE--- 406  
QY 189 GKASAKQRLKCA-----SLQKFGERAFKA-----WAVARLSQR 222  
Db 407 -----GVLOILICLOASAASNPFFSQAMDVLVOEFIQHGHFLFETAVILQMEWLVR--DG 460  
QY 223 FP-----KAEFAVSKLVDTLTKVHTE----- 244  
Db 461 VPPEASGHLKALINNMKIMSTVKVKBQHLHSMCTKRHRRCBYSHFMHHRDLGILL 520  
QY 245 ----- 244  
Db 521 VSAFNQVKNPFEETADGDVYPERCCCCIAVAHQCLRLQLQAASLSSTCVQILSGVQNI 580  
QY 245 ---CHGD-----LLE-----C 253  
Db 581 GICCCMDPKSVIPLHLHAFKLPALKSCQOHLNKLILDLQGAIEPQKTKAACNIC 640  
QY 254 ADDRADIK-----YICENQDSIS----- 273  
Db 641 TVDSQLAKLEETLOGSSYNAPFSGLSPSYRFQGIPLSSGSDLLWKWDALAYQNFV 700  
QY 274 -----KLKECCCKPILLEKS-----HC----- 289  
Db 701 FEEDRLQSVQIANHICSLIQGNVVQWKLNCIENPVLQRGVELAHHCQQLSISQAOTH 760  
QY 290 IAEVENDPAD-----LPSLAADFV----- 310  
Db 761 VCSHNNQCLPQVLOIYLKTLPTLLKSRVIRDLFLSCNGVNIIEINLYLDGIRNHSKAF 820  
QY 311 -----ESKOVCKNYAEAKVP 326  
Db 821 ETLIISLGEQQKSSIPGIDGLDEQKELSSNVGLYNQAYSDQSLSKPYARKLQAY 880  
QY 327 -----LGMFLYEYARRHPDYSVVLLRLAKTYETTLEK-----C 361  
Db 881 PKKRSVVQDIHISTINFL-----CVAFCLVSKAEASDESANDSEDTSGYDS 929  
QY 362 AAADP-----HE-----CY 370  
Db 930 TASEPLQHLMPCFSLLESVLPSPRMHQAADVMSMCRWIYMLSPIFRKQFYRLGGFQVCH 989  
QY 371 AKVF-----DEFKPLVEE-----PQNLIKQNCELFEQIGYKQFN 405  
Db 990 KLIFMVIQDLFRNPEEQKREGDRTMENQDLNRIQSPETITVKED----- 1035  
QY 406 ALLVRYTKKVPQVSTPTLVEVGR-----NLGKVGSKCKCKHPEAKRMPCAE----- 450  
Db 1036 ---LLSLTVKI-----DPTPTLSLKSADSLSKLESEHLSINVEQIPAVEAVPEETKVF 1089  
QY 451 -----DYLGSVLN 458  
Db 1090 MSRESETLQIGIRLLEALLAICLHGTRASQOKLELPLNQNLVETILLEMRDHL----- 1145  
QY 459 QLCVLHEKTPVSD-----RV----- 473  
Db 1146 -----KSKVTETELAKPLFDALLRVALGNHSADEHDDAMTEKSHQSEELSQQGDF 1198  
QY 474 -----TKCCTESLVNRRPCFSALVEDETY-----VPKEFNAETF--- 507  
Db 1199 SEEAEDSQCCS-----FKLLVEEGYEADSENPEDESETWDDGVDLKPFAESFIAS 1249  
QY 508 -----TFHADICTIS-----EKERQI----- 523

Db 1250 SSPNDLLENLSSQGBIIYIPEICTLEINLLSTGKAKLDVLAHVFESEFLKIIRQKKNIFLLM 1309  
QY 524 -----KKQTALVELVXHKPKA 539  
Db 1310 QQGTVKNLGGFLSILTQDSDFOACQORVLDLLVSLMSSRTCSBELTLLRIFLEKSPC 1369  
QY 540 TKEQLKAVM-----DDFAAF-----VEKC----- 558  
Db 1370 TEILLGLIKLIVESDITWSPSQYLTFFLLHTPNLSNGVSSQKCPGILNSKAMGLLRARV 1429  
QY 559 -----CKADDKETCFAE 570  
Db 1430 SOSKIEGSESPQOQLSSWHIAPHVLPPLQOPHSEGSISLWNVFECIHEPEST--TE 1487  
QY 571 EGKK-----LVAASQAALGLMSRLEVPCHALPOG-----LSPQOVIIIRGLV----- 614  
Db 1488 KGKTRKRNKSLVLLDSSFDGTENNRLE-----GAAVYVNGERLIEGCVHMISLG 1538  
QY 615 -----LOEP 618  
Db 1539 SKALIQWADPHTGTFTIFRVCMDSNDDTKVLLAQVESQENIFLPSKWQHLVLTVLOQ 1598  
QY 619 K-----HFTVSLRDQAAHAPVTLRASFADRTLQWISRWGQKKLISAPFLFYQRFEEVLL 673  
Db 1599 QGKKNHKGKISI-----WIS--GQRK-----P-----DVTL 1622  
QY 674 LF-----QEGGLKLA-----LNGOGLGATSMNQQA 698  
Db 1623 DFMLPRKTSLSSDSNKTFMIGHCLSSQESLQLAGKWDLGNLLFNPAKIG--SQEA 1678  
QY 699 L-----EQLREL-----RISG 709  
Db 1679 FYLYACGNHTSIMPKYKGPVNDYSKYINTEILOCEQIRELFMTSKDVIDIGLLIESLV 1738  
QY 710 SVQLYC 715  
Db 1739 VYTTYC 1744

Search completed: October 17, 2005, 08:36:37  
Job time : 53 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2005, 08:25:57 ; Search time 178 Seconds  
(without alignments)  
2065.578 Million cell updates/sec

Title: US-10-933-523-18

Perfect score: 3785

Sequence: 1 DAHKEVAHFKDLGLENFK.....leqirelriagsvqlcyvhs 718

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.1

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	82.0	609	1	ALBU_HUMAN
2	3079	81.3	609	2	Q68DN5
3	2942	77.7	600	1	ALBU_MACMU
4	2620	69.2	608	1	ALBU_FELCA
5	2614	69.1	584	2	Q7YSG3
6	2562	67.7	608	1	ALBU_CANFA
7	2504	66.2	608	2	Q95VB7
8	2475.9	65.4	607	1	ALBU_HORSE
9	2457	64.9	608	1	ALBU_RABIT
10	2450.9	64.8	607	1	ALBU_BOVIN
11	2432.9	64.3	607	1	ALBU_SHEEP
12	2426	64.1	608	1	ALBU_RAT
13	2416.8	63.9	605	1	ALBU_PIG
14	2409.8	63.7	607	2	Q68NH7
15	2387	63.1	608	2	Q6WDN9
16	2387	63.1	609	1	ALBU_MERUN
17	2379.9	62.9	583	2	Q6B3Z0
18	2378	62.8	608	1	ALBU_MOUSE
19	2374	62.7	608	2	Q8C7H3
20	2336	61.7	576	2	Q8C7C7
21	2045.8	54.1	417	2	Q85VG0
22	1950.7	51.5	396	2	Q81UK7
23	1557.9	41.2	615	1	ALBU_CHICK
24	1296.7	34.3	527	2	Q8JIA9
25	1256.7	33.2	609	1	FETA_PANTR
26	1252.7	33.1	609	1	FETA_HUMAN
27	1249.3	33.0	626	2	Q8UW05
28	1245.7	32.9	609	1	FETA_GORGO
29	1242.8	32.8	609	2	Q8MUJ5
30	1219.7	32.2	610	2	Q8MJ76
31	1206.6	31.9	607	1	ALB2_XENLA

32	1201.2	31.7	609	1	FETA_HORSE
33	1200.6	31.7	607	2	Q642F7
34	1183.9	31.3	608	2	Q7TSE3
35	1168.1	30.9	606	1	ALB1_XENLA
36	1096.8	29.0	624	2	Q8UW06
37	1086.8	28.7	605	1	FETA_MOUSE
38	1084.8	28.7	605	2	Q8BK65
39	1078.8	28.5	605	2	Q8BK56
40	1071.6	28.3	611	1	FETA_RAT
41	1059	28.0	599	1	AFAM_HUMAN
42	1050.7	27.8	400	2	Q8JIA7
43	961.4	25.4	603	2	Q9YGH6
44	948.7	25.1	611	1	AFAM_MOUSE
45	933.4	24.7	614	2	Q91134
46	932	24.6	608	1	AFAM_RAT
47	894.6	23.6	406	2	Q8JIA8
48	771	20.4	167	2	Q6UXK4
49	765.3	20.2	608	1	ALB1_SALSA
50	761.3	20.1	608	1	ALB2_SALSA
51	739	19.5	205	2	Q8CG74
52	716.9	18.9	417	2	Q8R0J9
53	709.5	18.7	336	1	LEGC_HUMAN
54	704.9	18.6	296	2	Q9NZ03
55	700.4	18.5	382	1	ALBU_RANCA
56	601.7	15.9	314	1	LEGC_MOUSE
57	531.4	14.0	1423	1	ALBU_PETNA
58	527.7	13.9	383	2	Q8C3K0
59	421	11.1	476	1	VTDB_RAT
60	421	11.1	476	2	Q68FY4
61	417.5	11.0	474	1	VTDB_HUMAN
62	415.2	11.0	551	2	Q42279
63	414.8	11.0	476	1	VTDB_RABIT
64	414	10.9	476	1	VTDB_MOUSE
65	408.5	10.8	484	2	Q9W6F5
66	375.3	9.9	3207	2	Q8MW03
67	359.7	9.5	3287	2	Q8T326
68	347.4	9.2	3542	2	Q9U5M2
69	344.9	9.1	3279	2	Q9N4B9
70	344.9	9.1	3843	2	Q9VU94
71	339.8	9.0	464	2	Q6DGV8
72	338.9	9.0	3843	2	Q9U5D0
73	335	8.9	1797	2	Q75X61
74	333.2	8.8	2002	2	Q75XH1
75	332.1	8.8	3796	2	Q97699
76	331.6	8.8	3666	2	Q6UDX0
77	331.5	8.8	3795	2	Q9TTK4
78	331.1	8.7	4283	2	Q8UWL7
79	331	8.7	122	2	Q90WZ8
80	329.6	8.7	1927	2	Q25262
81	329	8.7	1802	2	Q75X05
82	328.7	8.7	1820	2	Q75XM5
83	328.6	8.7	1813	2	Q75XE3
84	328.2	8.7	2316	2	Q8WXA6
85	327.2	8.6	1819	2	Q9ZLV0
86	326.9	8.6	2858	2	Q8IAK2
87	326.8	8.6	2443	2	Q96J17
88	326.6	8.6	2994	2	Q81I24
89	325	8.6	1927	2	Q75WX9
90	324.9	8.6	2118	1	ESP1_MOUSE
91	324.7	8.6	1894	2	Q75XB6
92	324.5	8.6	3954	2	Q6LEV2
93	324.3	8.6	1795	2	Q75XJ8
94	324	8.6	2154	2	Q9ZVC8
95	323.5	8.5	2342	2	Q8NG31
96	323.3	8.5	6885	1	SNE2_HUMAN
97	322.9	8.5	2553	2	Q6PUE8
98	322.2	8.5	3175	2	Q7QPT9
99	322.8	8.5	2418	1	SPCA_HUMAN
100	322.5	8.5	3346	2	Q8IEU9

ALIGNMENTS

RESULT 1  
ALBU\_HUMAN STANDARD; PRT; 609 AA.  
AC P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Serum albumin precursor.  
GN Name=ALB;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196112; PubMed=3009475;  
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,  
RA Beattie W.G., Dugaiczky A.;  
RT "Molecular structure of the human albumin gene is revealed by  
RT nucleotide sequence within q11-22 of chromosome 4.";  
RL J. Biol. Chem. 261:6747-6757(1986).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.  
RX MEDLINE=62081882; PubMed=6171778;  
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
RA Najarian R.C., Seeburg P.H., Wion K.L.;  
RT "The sequence of human serum albumin cDNA and its expression in E.  
RT coli.";  
RL Nucleic Acids Res. 9:6103-6114(1981).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.  
RX MEDLINE=82105994; PubMed=6275391;  
RA Dugaiczky A., Law S.W., Dennison O.E.;  
RT "Nucleotide sequence and the encoded amino acids of human serum  
RT albumin mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).  
RC TISSUE=Fetal liver;  
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,  
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;  
RT "Functional prediction of the coding sequences of 121 new genes  
RT deduced by analysis of cDNA clones from human fetal liver.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.  
RA Huang M.C., Wu H.T.;  
RT "The cDNA sequences of human serum albumin.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins K.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs K.A.,  
RA Fayeton J., Heiton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RP SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.Q.;  
RL (in) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RP SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Farrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RT regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RP SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RT acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RP SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RT 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RT affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).  
RN [16]  
RP VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
RT domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RT Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]

Mon Oct 17 12:55:14 2005

```
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huse K., Ishioka N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese alloalbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
[19]
RN VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
[20]
RN DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
human serum albumin whose precursor has an aberrant signal peptidase
cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
[21]
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Minchiotti L., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
[22]
RN VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
[23]
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Anaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
[24]
RN VARIANT CASEBOOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851; DOI=10.1016/0925-4439(91)90023-3;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
albumin: albumin Casebrook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
[25]
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703; DOI=10.1016/0167-4838(92)90207-T;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two alloalbumins with identical electrophoretic mobility are produced
Query Match 82.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 2.9e-70;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
DB 145 DVNCTAFHNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 204
```

```
QY 181 KLDLRLDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKABFAEYSLVTDLTk 240
DB 205 KLDLRLDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKABFAEYSLVTDLTk 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYEARHHPDYSVLLRLAKTYETTLKBC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYEARHHPDYSVLLRLAKTYETTLKBC 384
QY 361 CAAADPHCYAKVDFDEPKPLVEEPQNLIKQNCLEFEQQLGEYKFNALAVRYTKKVPQYST 420
DB 385 CAAADPHCYAKVDFDEPKPLVEEPQNLIKQNCLEFEQQLGEYKFNALAVRYTKKVPQYST 444
QY 421 PTLVEVSRLNKGKSGCKCKHPEAKRMPCAEADYLSVNLQNLVHLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNKGKSGCKCKHPEAKRMPCAEADYLSVNLQNLVHLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALAEVDYVYVPEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHKPAT 540
DB 505 LVNRRPCFSALAEVDYVYVPEFNAETFTFHADICTLSEKEROIKKQTALVELVKKHKPAT 564
QY 541 KEQLKAVNMDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVNMDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 2
Q68DN5 PRELIMINARY; PRT; 609 AA.
ID AC Q68DN5;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp779N1935.
GN Name=DKFZp779N1935;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RG The German cDNA Consortium;
RA Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; CR749331; CAH18185.1;
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PRO0803; AFETOPROTEIN.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Hypothetical protein.
SQ SEQUENCE 609 AA; 69402 MW; 3BA3AFF17BF99E94 CRC64;

Query Match 81.3%; Score 3079; DB 2; Length 609;
Best Local Similarity 99.1%; Pred. No. 1.2e-69;
Matches 580; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
```





```
Db 437 PTLVEVSNLKGKVGKACCKLPEAKRMPCAEDYLSVLNRLCVLHKTVPVSKVTKCCTES 496
Qy 481 LVNRRPCFSALEVDYTPYKFNATFTFHADICTLSEKRIQKOTATLVELVXHKPKAT 540
Db 497 LVNRRPCFSALELDYAVPKAFNATFTFHADICTLSEKRIQKOTATLVELVXHKPKAT 556
Qy 541 KEQLKAVNMDFAAFVEKCKCKADDKTCFAEGKKLVAAQAAL 583
Db 557 KEQLKGVNMDNFAAFVEKCKCKADDKTCFAEGKKLVAAQAAL 599

RESULT 4
ALBU_FELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN Name=ALB;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;
RA Hilger C., Grigioni F., Kohlen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X84842; CAA59279.1; -
CC PIR; JC4660; S57632.
CC
CC HSP; P02768; 1E7B.
CC
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; Serum albumin; 3.
CC PRINTS; P00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC
CC PROSITE; PS00212; ALBUMIN; 3.
CC Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
KW SIGNAL
FT 1 18 By similarity.
FT PROPEP 19 24 By similarity.
FT CHAIN 25 608 Serum albumin.
FT DOMAIN 25 205 Albumin 1.
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 27 27 Copper.
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
```

```
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;

Query Match 69.2%; Score 2620; DB 1; Length 608;
Best Local Similarity 82.0%; Pred. No. 5.8e-58;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

Qy 1 DAHSEVAHREFKDLGEENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 EAHQSEIAHRPNLDGEEHFRGLVAVFSQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLTRETGEMADCCAKSPERNECFLOHKDDNPRLVVRPEV 120
Db 85 NCEKSLHELLGDKLCTVASLDRKYGEMADCCAKSPERNECFLOHKDDNPFGQLVTPEA 144
Qy 121 DVMCTAHNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
Db 145 DAMCTAFHENSQRFGLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 204
Qy 181 KLDELDRSGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRFKAAEFAEVSCLVDTLTK 240
Db 205 KVDALREKVLASSAKERLKCASLQKFGGERAFKAWAVARLSQRFKAAEFAEISKLVTDLAK 264
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKKECKPPLLEKSHCIAEVENDEMPA 300
Db 265 IHKECCHGDLLECADRADLAKYICENQDSISSTKLKECCGKPVLEKSHCISEVERDELPA 324
Qy 301 DLPSLAADPVESKOVKNYAEAKOVFLCMVLEYEARHPDYSVLLLLAKTYETTLK 360
Db 325 DLPLAVDFVEDKEVKCNVQAKOVFLGTFLEYEYRRHPPEYSVLLLLAKTYETTLK 384
Qy 361 CAAADPHCYAKVDFEKPPLVEEPONLIKQCELFQELGEYKFNQALLVRYTKVPQVST 420
Db 385 CATDPPACYAHVDFEKPPLVEEPHNLVKTNCLEFELGEGFNALLVRYTKVPQVST 444
Qy 421 PTLVEVSNLKGKVGKACCKLPEAKRMPCAEDYLSVLNRLCVLHKTVPVSKVTKCCTES 480
Db 445 PTLVEVSNLKGKVGKACCKLPEAKRMPCAEDYLSVLNRLCVLHKTVPVSKVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYTPYKFNATFTFHADICTLSEKRIQKOTATLVELVXHKPKAT 540
Db 505 LVNRRPCFSALEVDYTPYKFNATFTFHADICTLPEAKRIQKQSLVLLXHKPKAT 564
Qy 541 KEQLKAVNMDFAAFVEKCKCKADDKTCFAEGKKLVAAQAAL 583
Db 565 KEQLKGVNMDNFAAFVEKCKCKADDKTCFAEGKKLVAAQAAL 607

RESULT 5
QYSG3
ID QYSG3 PRELIMINARY; PRT; 584 AA.
AC QYSG3;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Albumin (Pragmat).
GN Name=alb;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
```



```
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 27 27 Copper (By similarity).
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 340 385 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
FT CONFLICT 1 26 MKWTFISLFLFSSAYSGRLVRREA -> MDT (in Ref. 2).
FT CONFLICT 146 146 A -> R (in Ref. 2).
FT CONFLICT 206 206 I -> T (in Ref. 2).
FT CONFLICT 349 349 V -> A (in Ref. 2).
FT CONFLICT 359 359 S -> A (in Ref. 1).
FT CONFLICT 448 448 V -> VV (in Ref. 5).
FT CONFLICT 474 474 D -> E (in Ref. 1).
FT CONFLICT 608 AA; 68606 MW; 3CF1C9FF7DD8FC06 CRC64;
SQ SEQUENCE 608 AA; 68606 MW; 3CF1C9FF7DD8FC06 CRC64;

Query Match 67.7%; Score 2562; DB 1; Length 608;
Best Local Similarity 79.8%; Pred. No. 1.7e-56;
Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDLGEENFKALVITAPQYLOQCPPEHDHVLNVETSEFAKTCVADESAE 60
DB 25 EAYKSEIARHYNDGEEHPRGLVAVFAQYLOQCPPEHDHVLNVETSEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPINLPRIVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPINLPRIVRPEV 144

QY 121 DVMCTAFHNDNEFTLKKLYEYIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
DB 145 DALCAAFQDNEQLFLGKLYEYIARRHPYFYAPPELLFYAQYKGVFAECCQAADKAACLLP 204

QY 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 205 KTEALREKVLSSAKERPKCASLOKFGGERAFKAWAVARLSQRPKADFAEISKVVTDLTK 264

QY 241 VHECCHGDLLECADDRLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300
DB 265 VHECCHGDLLECADDRLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYTETLEKC 360
DB 325 DLPSLAADFVEDKEVKCKNYAEAKDVFLGTFLEYARRHPDYSVVLLRLAKTYEATLEKC 384

QY 361 CAADPHCYAKVDFEFPKLPVEEPQNLIKONCELFEOIGYKFNALLVRYTKYKQPVST 420
DB 385 CATDDPPPTCYAKVDFEFPKLPVEEPQNLVKTNCLEFELGEYGFONALLVRYTKQAPVST 444

QY 421 PTLVEVSRNLGVKSGKCKKHPRKMPCAEDVLSVVLNQLCVLHKTTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVKSGKCKKHPRKMPCAEDVLSVVLNQLCVLHKTTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQATLVELVKGHPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQATLVELVKGHPKAT 564

QY 541 KEQLKAVMDDFAAVFVEKCKCADDKETCFABEGKLVAAASQAAL 583
DB 565 DEQLTKVMGDFGAFVEKCKCADDKETCFABEGKLVAAASQAAL 607
```

```
RESULT 7
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_taxid=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RL Submitted (SRP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP418550; AAL08579.1; -.
DR HSP; P02768; IHK1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; ESEABB28E1C66E54 CRC64;

Query Match 66.2%; Score 2504; DB 2; Length 608;
Best Local Similarity 76.3%; Pred. No. 5.2e-55;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDLGEENFKALVITAPQYLOQCPPEHDHVLNVETSEFAKTCVADESAE 60
DB 25 DAHKEVAHRPKDLGEENFKALVITAPQYLOQCPPEHDHVLNVETSEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPINLPRIVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPINLPRIVRPEV 144

QY 121 DVMCTAFHNDNEFTLKKLYEYIARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
DB 145 EAMCTSFQENAVTFMGHLYEVARHPYFYAPPELLFYAEKYSAIMTECCGEADKAACITP 204

QY 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPKADFAEISKVVTDLTK 264

QY 241 VHECCHGDLLECADDRLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300
DB 265 VHECCHGDLLECADDRLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYTETLEKC 360
DB 325 DLPSLAADFVEDKEVKCKNYAEAKDVFLGTFLEYARRHPDYSVVLLRLAKTYEATLEKC 384

QY 361 CAADPHCYAKVDFEFPKLPVEEPQNLIKONCELFEOIGYKFNALLVRYTKYKQPVST 420
DB 385 CAADPHCYAKVDFEFPKLPVEEPQNLVKTNCLEFELGEYGFONALLVRYTKQAPVST 444

QY 421 PTLVEVSRNLGVKSGKCKKHPRKMPCAEDVLSVVLNQLCVLHKTTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVKSGKCKKHPRKMPCAEDVLSVVLNQLCVLHKTTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQATLVELVKGHPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKFNAAETFTFHADICTLSEKERQIKKQATLVELVKGHPKAT 564

QY 541 KEQLKAVMDDFAAVFVEKCKCADDKETCFABEGKLVAAASQAAL 583
```



[2]  
 RN REVISIONS TO 322-323 AND 506-507.  
 RA Sheffield W.P.;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U18344; AAB58347.2; -.  
 DR HSPF; P02768; 1E7B.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN\_3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
 FT SIGNAL 1 18 By similarity.  
 FT PROPEP 19 24 By similarity.  
 FT CHAIN 25 608 Serum albumin.  
 FT DOMAIN 25 205 Albumin 1.  
 FT DOMAIN 212 397 Albumin 2.  
 FT DOMAIN 404 595 Albumin 3.  
 FT METAL 27 27 Copper.  
 FT DISULFID 77 86 By similarity.  
 FT DISULFID 99 115 By similarity.  
 FT DISULFID 114 125 By similarity.  
 FT DISULFID 148 193 By similarity.  
 FT DISULFID 192 201 By similarity.  
 FT DISULFID 224 270 By similarity.  
 FT DISULFID 269 277 By similarity.  
 FT DISULFID 289 303 By similarity.  
 FT DISULFID 302 313 By similarity.  
 FT DISULFID 340 385 By similarity.  
 FT DISULFID 384 393 By similarity.  
 FT DISULFID 416 462 By similarity.  
 FT DISULFID 461 472 By similarity.  
 FT DISULFID 485 501 By similarity.  
 FT DISULFID 500 511 By similarity.  
 FT DISULFID 538 583 By similarity.  
 FT DISULFID 582 591 By similarity.  
 SQ SEQUENCE 608 AA; 68909 MW; 9ECAFDA86B1EF09 CRC64;  
 Query Match 64.9%; Score 2457; DB 1; Length 608;  
 Best Local Similarity 74.3%; Pred. No. 8.2e-54;  
 Matches 434; Conservative 78; Mismatches 72; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFDKLGSENFKALVIAFAQYLOQCPEFDHVLVNEVTEFAKTCVADESAAE 60  
 DB 25 EAHKSEIAHRFNDVGEHEHIGLVITFSQYLQKCPYEEHAKLVKEVTDLAKACVADESAA 84  
 QY 61 NCDKSLHITLFGDKLCTVATLRTYTGEMADCCAKQEPERNECFLOHKDDNPRLVVRPE 120  
 DB 85 NCDKSLHIDIFGDKICALPSLRTYGDVADCCCKEPERNECFLLHHKDDKPDLPFPARPEA 144  
 QY 121 DVMCTAFHNEBTFLKKLYETARRHPYFAPPELLFPFAKRYKAAFTCCQAAADKACLLP 180  
 DB 145 DVLCKAFHDDKAFGHLYEVARHPYFAPPELLYIAQYKAILTECCAAADKAGCLTP 204  
 QY 181 KLDELREDEKASSAKQRLKCAISLQKFGERAFAKAWAVARLSQRFPAEFAEVSKLVTDLTK 240

DB 205 KLDALGKSLISAQERLCASIOKFGDRAYKAWALVRLSORFPKADFTDISKIYDITK 264  
 QY 241 VHTCCCHGDLLEACADRADLAKYICENQDSISSKLKECCPKLEKSHCIAEVENDEMPA 300  
 DB 265 VHKECCHGDLLEACADRADLAKYICENQDSISSHLKECCDKPILKHAHCYGLHNDETPA 324  
 QY 301 DLPSLAADFVSKDVCNKYAEAKOVFLGMFLYEVARRHPDYSVLLLLBLAKTYETTLKCC 360  
 DB 325 GLPAVABEFVEDKDVCKNYEBAKDLFLGKFLYESSRRHPDYSVLLLLBLGKAYEATLKCC 384  
 QY 361 CAADPHCEYAKVFDEKPLVEEPQNLKONCELFQOLGEYKFNQALLVRYTKKVPQVST 420  
 DB 385 CATDDPHACYAKVLDEFPQPLVDEPKNLVKNCLEYEQDYNFNQALLVRYTKKVPQVST 444  
 QY 421 PTLVEVSRLNKGKVGSKCKHPKPEAKMPCAEYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480  
 DB 445 PTLVEISRLNKGKVGSKCKHPKPEARLPCVEDYLSVVLNRLCVLHKEKTPVSEKVTCCSES 504  
 QY 481 LVNRRPCPSALVEDYTYVPKSFNAETFTFHADICTLSEKEROIKKQATLVELVHKHPAT 540  
 DB 505 LVDRRPCFSALGPDETYYPKSFNAETFTFHADICTLPETERKIKKQATLVELVHKHPAT 564  
 QY 541 KEQLKAYMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALG 584  
 DB 565 NDQLKTVVGEFTALLDKCCSAEDKEACFAVEGPKLVESKATLG 608  
 RESULT 10  
 ALBU\_BOVIN  
 ID ALBU\_BOVIN STANDARD; PRT; 607 AA.  
 AC P02769; O02787;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-PEB-1996 (Rel. 33, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Serum albumin precursor (Allergen Bos d 6) (BSA).  
 GN Name=ALB;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;  
 RT "Bovine serum albumin: cDNA sequence and expression.";  
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RC TISSUE=Liver;  
 RA Barry T., Power S., Gannon F.;  
 RT "The bovine serum albumin mRNA.";  
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=21195144; PubMed=11298124;  
 RA Hilger C., Grignon F., De Beaufort C., Michel G., Freilinger J.,  
 RA Hengges F.;  
 RT "Differential binding of IgG and IgA antibodies to antigenic  
 RT determinants of bovine serum albumin.";  
 RL Clin. Exp. Immunol. 123:387-394 (2001).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RA Wu H.T., Huang M.C.;  
 RT "The complete cDNA sequence of bovine serum albumin.";  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-32.  
 RX MEDLINE=80024278; PubMed=488109;  
 RA McGilivray R.T.A., Chung D.W., Davie E.W.;  
 RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-  
 RT terminal sequence of preproalbumin.";

RL Eur. J. Biochem. 98:477-485 (1979).  
RN [6]  
RP SEQUENCE OF 19-28.  
RX MEDLINE=77134075; PubMed=843354;  
RA Patterson J.E., Geller D.M.;  
RT "Bovine microsomal albumin: amino terminal sequence of bovine  
RT proalbumin.";   
RL Biochem. Biophys. Res. Commun. 74:1220-1226 (1977).  
RN [7]  
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
RX MEDLINE=91083649; PubMed=2260975;  
RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;  
RT "Rapid confirmation and revision of the primary structure of bovine  
RT serum albumin by ESIMS and Frit-FAB LC/MS.";   
RL Biochem. Biophys. Res. Commun. 173:639-646 (1990).  
RN [8]  
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.  
RA Brown J.R.;  
RT "Structure of bovine serum albumin.";   
RL Fed. Proc. 34:591-591 (1975).  
RN [9]  
RP REVISIONS TO 190-195.  
RA Brown J.R.;  
RT Submitted (APR-1975) to the PIR data bank.  
RN [10]  
RP SEQUENCE OF 25-64.  
RX PubMed=2379503;  
RA Strawich E., Glimcher M.J.;  
RT "Tooth 'enamelins' identified mainly as serum proteins. Major  
RT 'enamelin' is albumin.";   
RL Eur. J. Biochem. 191:47-56 (1990).  
RN [11]  
RP SEQUENCE OF 25-41.  
RX MEDLINE=88267456; PubMed=3389500;  
RA Hsieh J.C., Lin F.P., Tam M.F.;  
RT "Electroblotting onto glass-fiber filter from an analytical  
RT isoelectrofocusing gel: a preparative method for isolating proteins  
RT for N-terminal microsequencing.";   
RL Anal. Biochem. 170:1-8 (1988).  
RN [12]  
RP SEQUENCE OF 163-172.  
RX PubMed=2474609;  
RA Carraway R.E., Cochran D.E., Boucher W., Mitra S.P.;  
RT "Structures of histamine-releasing peptides formed by the action of  
RT acid proteases on mammalian albumin(s).";   
RL J. Immunol. 143:1680-1684 (1989).  
RN [13]  
RP SEQUENCE OF 402-433.  
RX MEDLINE=8203364; PubMed=7283978;  
RA Reed R.G., Putnam F.W., Peters T. Jr.;  
RT "Sequence of residues 400-403 of bovine serum albumin.";   
RL Biochem. J. 191:867-868 (1980).  
RN [14]  
RP SEQUENCE OF 437-451.  
RA Vilbois F.;  
RL Submitted (AUG-1998) to Swiss-Prot.  
RN [15]  
RP DISULFIDE BONDS.  
RA Brown J.R.;  
RT "Structure of serum albumin: disulfide bridges.";   
RL Fed. Proc. 33:1389-1389 (1974).  
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Plasma.  
CC -1- ALLERGEN: Causes an allergic reaction in human.  
CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -1- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M73993; AAA51411.1; -;  
DR EMBL; X58989; CAA41735.1; -;  
DR EMBL; Y17769; CAA76847.1; -;  
DR EMBL; AF542068; AAN17824.1; -;  
DR PIR; A38885; ABBOS.  
DR HSSP; P02768; IHKL.  
DR InterPro; IPR001703; Alphafetoprot.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; Serum\_albumin; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Allergen; Copper; Direct protein sequencing; Lipid-binding;  
KW Metal-binding; Polymorphism; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 607 Serum albumin.  
FT DOMAIN 25 204 Albumin 1.  
FT DOMAIN 211 396 Albumin 2.  
FT DOMAIN 403 594 Albumin 3.  
FT METAL 27 27 Copper (By similarity).  
FT DISULFID 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 192  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312  
FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
FT VARIANT 214 214 A -> T.  
FT CONFLICT 58 58 Missing (in Ref. 10).  
FT CONFLICT 302 302 C -> K (in Ref. 8).  
FT CONFLICT 304 305 KP -> PC (in Ref. 8).  
FT CONFLICT 324 324 N -> D (in Ref. 8).  
FT CONFLICT 394 395 ST -> TS (in Ref. 8).  
FT CONFLICT 437 437 K -> R (in Ref. 14).  
FT CONFLICT 493 494 SE -> ES (in Ref. 8).  
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
  
Query Match 64.8%; Score 2450.9; DB 1; Length 607;  
Best local Similarity 75.8%; Pred. No. 1.2e-53;  
Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;  
  
QY 1 DAUKSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPFDDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DTHKSEAHHRFKDGLGEHFKGLVLIAPSOYLQCCPFDDHVKLVNELTEFAKTCVADESAA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120  
DB 85 GCEKSLHTLFGDELCKVASLRETYGMDADCCQEPERNECFLSHKDDSPDLPKL-KPDP 143  
QY 121 DVMTAFHDNEETFLKYLVEIARRHFFYFAPELLFPKRYKAAPTECCQADKAACLLP 180  
DB 144 NTLGDEFKADKKFWGKYLVEIARRHFFYFAPELLFYANKYNGVVFQCCQADKGACLLP 203  
QY 181 KLDELREGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFKAEVSKLVTDLTK 240

```
Db 204 KIETMRKVLASSARORLRCASIQKFGERALKAWSVARLSQKFFKABFVETVLVTLTK 263
Qy 241 VHTCCGDLLEACADRADLAKYICENODSISSSKLKECCPELLEKSHCIAEVENDEMPA 300
Db 264 VHKCCGDLLEACADRADLAKYICDNQDTISSKLKECCDPLLEKSHCIAEVEKDAIPE 323
Qy 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYFYEARHPDYVSVLLRLAKTYETTLK 360
Db 324 NLPLTADFAEDKDVCKNYQEAQDAFLGSLFYYSRRHPEYAVSVLLRLAKEYEATLEEC 383
Qy 361 CAADPHCYAKVDFEFPLVBEONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 420
Db 384 CAKDPHACYATVDFKLHLVDEPNLIKQNCDOFEKLGEGYFQNALIVRYTKVPQVST 443
Qy 421 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVNLQNLVLEKTPVSDRVTKCCTES 480
Db 444 PTLVEVSRLGKVGTRCTCKPESERMPCTEDYLSLNLRLCVLHEKTPVSEKVKCCTES 503
Qy 481 LVNRRPCFSALVEDETYVPKFNABFTFTFHADICTLSEKEROIKQTALVELVKKPKAT 540
Db 504 LVNRRPCFSALTPDQTYVPKAFDEKLTFTFHADICTLPTDEKQIKQTALVELLKHKPKAT 563
Qy 541 KEOLKAVMDDDFAAFVEKCKCKADKCTCFABEGKLVAAASQAAL 583
Db 564 BEOLKTVMENFVAFVDDKCAADDKEACFAVBGPGLVSTQAL 606

RESULT 11
ALBU SHEEP
ID ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serum albumin precursor.
GN Name=ALB;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=9009888; PubMed=2602160;
RX Brown W.M., Driegielewska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL Nucleic Acids Res. 17:10495-10495(1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X17055; CAA34903.1; -
CC PIR; S06936; ABSHS.
CC HSSP; P02768; 1HK1.
CC InterPro; IPR001703; Alphafetoprot.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; Serum_albumin; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
```

```
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT SIGNAL 1 18 By similarity.
FT PROPEP 19 24 By similarity.
FT CHAIN 25 607 Serum albumin.
FT DOMAIN 25 204 Albumin 1.
FT DOMAIN 211 396 Albumin 2.
FT DOMAIN 403 594 Albumin 3.
FT METAL 27 27 Copper (By similarity).
FT DISULFID 77 86 By similarity.
FT DISULFID 99 115 By similarity.
FT DISULFID 114 125 By similarity.
FT DISULFID 147 192 By similarity.
FT DISULFID 191 200 By similarity.
FT DISULFID 223 269 By similarity.
FT DISULFID 268 276 By similarity.
FT DISULFID 288 302 By similarity.
FT DISULFID 301 312 By similarity.
FT DISULFID 339 384 By similarity.
FT DISULFID 383 392 By similarity.
FT DISULFID 415 461 By similarity.
FT DISULFID 460 471 By similarity.
FT DISULFID 484 500 By similarity.
FT DISULFID 499 510 By similarity.
FT DISULFID 537 582 By similarity.
FT DISULFID 581 590 By similarity.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 64.3%; Score 2432.9; DB 1; Length 607;
Best Local Similarity 75.0%; Pred. No. 3.4e-53;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

Qy 1 DAHKSEVAHREFKDLGEENFKALVIAFAQYLQCCPFEDHVLVNEVTFKTCVADESAE 60
Db 25 DTHKSEIAHREFNDLGEENFQGLVLIAPFQYLQCCPFEDHVLVNEVTFKTCVADESHA 84
Qy 61 NCDKSLHTLPGDKLCTVATLRETTGEMADCCAKQEPERNECFLOHKKDNDPNLRLVRP 120
Db 85 GCDKSLHTLFGDELCKVATLRETTGEMADCCCKQEPERNECFLNHKKDSDPLPKL-KPEP 143
Qy 121 DVMCTAFHDNETFLKYLVEIARHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 144 DTLCAEFKADKKFWGKLYEVARHPYFYAPELLYYANKYNGVFECCQADKGAACLLP 203
Qy 181 KLDELRLDEGKASSAKQRLKCSAQKFGGERAFKAMAVARLSQRPKAEFAEVSCLVTLTK 240
Db 204 KIDAREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKADFTDVTKIIVTLTK 263
Qy 241 VHTCCGDLLEACADRADLAKYICENODSISSSKLKECCPELLEKSHCIAEVENDEMPA 300
Db 264 VHKCCGDLLEACADRADLAKYICDNQDTISSKLKECCDPLLEKSHCIAEVEKDAVPE 323
Qy 301 DLPSLAADFVESKDVCKNYABAKOVFLGMFLYFYEARHPDYVSVLLRLAKTYETTLK 360
Db 324 NLPLTADFAEDKDVCKNYQEAQDAFLGSLFYYSRRHPEYAVSVLLRLAKEYEATLEDC 383
Qy 361 CAADPHCYAKVDFEFPLVBEONLIKONCELFEOQLGEYKFNALLVRYTKVPQVST 420
Db 384 CAKDPHACYATVDFKLHLVDEPNLIKQNCDOFEKLGEGYFQNALIVRYTKAPQVST 443
Qy 421 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVNLQNLVLEKTPVSDRVTKCCTES 480
Db 444 PTLVEVSRLGKVGTRCTCKPESERMPCTEDYLSLNLRLCVLHEKTPVSEKVKCCTES 503
Qy 481 LVNRRPCFSALVEDETYVPKFNABFTFTFHADICTLSEKEROIKQTALVELVKKPKAT 540
Db 504 LVNRRPCFSALTPDQTYVPKAFDEKLTFTFHADICTLPTDEKQIKQTALVELLKHKPKAT 563
Qy 541 KEOLKAVMDDDFAAFVEKCKCKADKCTCFABEGKLVAAASQAAL 583
Db 564 BEOLKTVMENFVAFVDDKCAADDKEGCFVLEGPGLVASTQAAL 606
```

RESULT 12  
ALBU\_RAT  
ID ALBU\_RAT STANDARD; PRT; 608 AA.  
AC P02770; P11382;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].  
GN Name=Alb;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=81223722; PubMed=7017712;  
RA Sargent T.D., Yang M., Bonner J.;  
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).  
[2]  
RN SEQUENCE OF 1-38, AND PROCESSING.  
RP MEDLINE=77249657; PubMed=893447;  
RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;  
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis of the direct translation product of albumin messenger RNA.";  
RL J. Biol. Chem. 252:6846-6855(1977).  
[3]  
RN SEQUENCE OF 25-222.  
RP MEDLINE=78109429; PubMed=564345;  
RA Isemura S., Ikenaka T.;  
RT "Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage of rat serum albumin.";  
RL J. Biochem. 83:35-48(1978).  
[4]  
RN SEQUENCE OF 223-288 AND 572-608.  
RP MEDLINE=76260153; PubMed=956149;  
RA Isemura S., Ikenaka T.;  
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequences of four fragments.";  
RL J. Biochem. 79:1183-1196(1976).  
[5]  
RN SEQUENCE OF 166-174.  
RP TISSUE=Plasma;  
RC MEDLINE=87194805; PubMed=2437111;  
RA Caraway R.E., Mitra S.P., Cochran D.E.;  
RT "Structure of a biologically active neurotensin-related peptide obtained from pepsin-treated albumin(s).";  
RL J. Biol. Chem. 262:5968-5973(1987).  
[6]  
RN COPPER-BINDING.  
RP MEDLINE=79001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
RL Cancer Res. 38:3483-3486(1978).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and blood flow (Potential).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC -----  
CC EMBL; V01222; CAA24532.1; --  
CC PIR; A93872; ABRTS.  
CC HSSP; P02768; IHKI.  
CC Rat-heart-2DPAGE; P02770; --  
CC RGD; 2085; Alb.  
CC InterPro; IPR001703; Alphafetoprot.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; Serum\_albumin; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.  
CC ProDom; PD002486; Serum albumin; 1.  
CC SMART; SM00103; ALBUMIN; 3.  
CC PROSITE; PS00212; ALBUMIN; 3.  
KW Copper; Direct protein sequencing; Lipid-binding; Metal-binding;  
KW Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 608 Serum albumin.  
FT PEPTIDE 166 174 Neurotensin-related peptide.  
FT DOMAIN 25 205 Albumin 1.  
FT DOMAIN 212 397 Albumin 2.  
FT DOMAIN 404 595 Albumin 3.  
FT METAL 27 27 Copper.  
FT DISULFID 77 86 By similarity.  
FT DISULFID 99 115 By similarity.  
FT DISULFID 114 125 By similarity.  
FT DISULFID 148 193 By similarity.  
FT DISULFID 192 201 By similarity.  
FT DISULFID 224 270 By similarity.  
FT DISULFID 269 277 By similarity.  
FT DISULFID 289 303 By similarity.  
FT DISULFID 302 313 By similarity.  
FT DISULFID 340 385 By similarity.  
FT DISULFID 384 393 By similarity.  
FT DISULFID 416 462 By similarity.  
FT DISULFID 461 472 By similarity.  
FT DISULFID 485 501 By similarity.  
FT DISULFID 500 511 By similarity.  
FT DISULFID 538 583 By similarity.  
FT DISULFID 582 591 By similarity.  
FT VARIANT 262 262 V -> L.  
FT CONFLICT 174 174 Y -> L (in Ref. 5).  
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;

Query Match 64.1%; Score 2426; DB 1; Length 608;  
Best Local Similarity 73.4%; Pred. No. 5.1e-53;  
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;  
Qy 1 DAKSEVAHRPKDILGEENFKALVLI AFAQYLQOCPPFDHVKLVNEVTEFAKTCVADERSAE 60  
Db 25 EAHKSEFAHRPKDILGEQHFGLVLI AFSQYLQKCPYEEHKLVOEVTDFAKTCVADENAE 84  
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVREPV 120  
Db 85 NCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKQEPERNECFLOHKDDNPNLPLFPQPEA 144  
Qy 121 DVMTATHDNEETFLKKYLVEIARRHPYFYFAPELLFAKRYKAAFTCCQAADKAACLLP 180  
Db 145 EAMCTSPQENPTSLFGLHYLHEVARRHPYFYFAPELLYAAKYNEVLTOCCTESDAAACLLP 204  
Qy 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAFVSKVLVTLTK 240  
Db 205 KLDVKEKALVAVRQRMKSSMORFGERAFKAWAVARMSORFFNASEFAEITKLATDVK 264  
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDENPA 300  
Db 265 INKECCHGDLLECCADRAELAKYWCENQATISSKLQACCDKPVLOKSOCLAEETHDNLPA 324  
Qy 301 DLPSLAADFVSKDKVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETLEKC 360  
Db 325 DLPSIAADFVEDKEVCKNYAEAKDVFGLTFLYEYSRRHPDYSVSLLRLAKKYEATLEKC 384





```

25-OCT-2004 (TrEMBLrel. 28, Created)
25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Albumin.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Sun S., Deng J., Zhou Y., Lu J., Wu X.;
RT "Porcine serum albumin gene."
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY663543; AAT98610.1; -
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; Serum_albumin; 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 3.
DR SEQUENCE 607 AA; 69691 MW; 7B8DAL3543CA99D8 CRC64;

Query Match 63.7%; Score 2409.8; DB 2; Length 607;
Best Local Similarity 71.7%; Pred. No. 1.3e-52;
Matches 440; Conservative 70; Mismatches 72; Indels 32; Gaps 2;

Qy 1 DAHSEVAHREPKDGEENFKALVLIAPAOYLQOCPPFDHVKLVNEVTEFAKTCVADESA 60
D 25 DTYKSEIAHREPKDGEQYFGKGLVLIAPSOHLQOCPPYEEHVKLVNEVTEFAKTCVADESA 84
D 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLRLVPEV 120
D 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLRLVPEV 143
D 144 VALCADFOEDQKFWGKYLVEIARRHPYFYAPPELLYAIYKOVFSECCQAADKAAACLLP 203
Qy 121 DVMTCTAHFNDNEFLKYLVEIARRHPYFYAPPELLYAIYKOVFSECCQAADKAAACLLP 180
D 204 KIEHLREKVLTSAAKORLKASIQKFGERAFAKAWARLSORFPKADFTSEIKVITDLAK 263
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDTSISKLKCECKEPLLEKSHCIAEVENDEMPA 300
D 264 VHKCCCHGDLLECCADRADLAKYICENQDTSISKLKCECKEPLLEKSHCIAEVENDEMPA 323
Qy 301 DLPDLAADFVSKDVCNRYAEAKDVFGLGMFLYVYARRHPDYVSVLLRLAKTYETTTLEK 360
D 324 DLNPLEHDFVEDKEVCNRYAEAKDVFGLGMFLYVYARRHPDYVSVLLRLAKTYETTTLEK 383
Qy 361 CAADPHECYAKVDFEKPVEEPQNLIKONCELFQOLGEYKFNALIVRYTKKVPQVST 420
D 384 CAKEDPPACYATVDFKQPLVDEPNLIKONCELFQOLGEYKFNALIVRYTKKVPQVST 443
Qy 421 PTLVEVSRNLGKVGSKCKCKHPAEKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCCTES 480
D 444 PTLVEVARKLGLVGRCCRPDEERLSCADYLSVLNQLCVLHEKTPVSEKVTYKCCCTES 503
Qy 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540
D 504 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 563
Qy 541 KEQLKAVMDFAAEVKECKADDEKTCFAEKGKLVLAASQAALGIMSPRLEVPCSHALPQ 600
D 564 BEQLRTVLGNFAAFVQKCAAPDHACFAVEGPKF----- 598
Qy 601 GLSPQGVIIIRGLV 614
D 599 -----VIEIRGIL 606
```

RESULT 15  
Q6WDN9

```

ID Q6WDN9 PRELIMINARY; PRT; 608 AA.
AC Q6WDN9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Preproalbumin precursor.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Mihaylova-Todorova S., Choe S.M., Westfall D.P.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY294645; AAQ20088.1; -
DR HSP; P02768; IAO6.
DR DR GO:0005615; C:extracellular space; IEA.
DR DR GO:0005386; F:carrier activity; IEA.
DR DR GO:0006810; P:transport; IEA.
DR InterPro: IPR001703; Alphafetoprot.
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; Serum_albumin; 3.
DR PRINTS: PR00803; AFETOPROTEIN.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 3.
DR Signal.
KW SIGNAL.
FT CHAIN 1 21 Potential.
FT SIGNAL 25 608 albumin.
SQ SEQUENCE 608 AA; 69889 MW; BBD510A78D0261BA CRC64;

Query Match 63.1%; Score 2387; DB 2; Length 608;
Best Local Similarity 72.2%; Pred. No. 5e-52;
Matches 421; Conservative 87; Mismatches 75; Indels 0; Gaps 0;

Qy 1 DAHSEVAHREPKDGEENFKALVLIAPAOYLQOCPPFDHVKLVNEVTEFAKTCVADESA 60
D 25 EAHKSEIAHREPKDGEHFGKGLVLIAPSOHLQOCPPFEHVKLVNEVTEFAKTCVADESA 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLRLVPEV 120
D 85 NCGKAIATLFGDKVCAIPSLRETYGEMADCCAKOEPRNECFLOHKDDNPNLRLVPEV 144
Qy 121 DVMTCTAHFNDNEFLKYLVEIARRHPYFYAPPELLYAIYKOVFSECCQAADKAAACLLP 180
D 145 EALCTAFKNNDRPIGHLYEVSRRHPYFYAPPELLYAIYKOVFSECCQAADKAAACLLP 204
Qy 181 KLDELDEGKASSAKORLKASIQKFGERAFAKAWARLSORFPKADFTSEIKVITDLTK 240
D 205 KLDAIKEALVSSAQORLKASIQKFGERAFAKAWARLSORFPKADFTSEIKVITDLTK 264
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDTSISKLKCECKEPLLEKSHCIAEVENDEMPA 300
D 265 VTKCCCHGDLLECCADRADLAKYICENQDTSISKLKCECKEPLLEKSHCIAEVENDEMPA 324
Qy 301 DLPDLAADFVSKDVCNRYAEAKDVFGLGMFLYVYARRHPDYVSVLLRLAKTYETTTLEK 360
D 325 ELPLDVAFDVEDKEVCNRYAEAKDVFGLGMFLYVYARRHPDYVSVLLRLAKTYETTTLEK 384
Qy 361 CAADPHECYAKVDFEKPVEEPQNLIKONCELFQOLGEYKFNALIVRYTKKVPQVST 420
D 385 CAEADPHACYAKVDFEKLPLIDEPKLQVQNCFLDKLGEYGFONALVRYTKQAPVST 444
Qy 421 PTLVEVSRNLGKVGSKCKCKHPAEKRMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCCTES 480
D 445 PTLVEYARKLGLSVGTCCKSLPETERLSCTENYALILNRLCILHEKTPVSDRVTKCCCTES 504
Qy 481 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540
D 505 LVNRRPCFSALEVDVETVYKPEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 564
```

QY 541 KEQLKAVNDFAAFVEKCKKADKCTCFABEGKLVAAASQAAL 583  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 565 BEQMTVMGDFAAFLKCCDADNKEACFTEDGPKLVAKQATL 607

## RESULT 16

ALBU MERUN  
ID ALBU MERUN STANDARD; PRT; 609 AA.  
AC Q35090;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Serum albumin precursor.  
GN Name=ALB;  
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
OC Meriones.  
OX NCBI\_TaxID=10047;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=MGS IDR; TISSUE=Liver;  
-RX MEDLINE=9811663; PubMed=9455485;  
RA Yoshida K., Seto-Onshima A., Sinochara H.;  
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
RL DNA Res. 4:351-354(1997).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB006197; BAA21765.1; -;  
CC DR PIR; JCS838; JCS838.  
CC DR HSP; P02768; IE7B.  
CC DR InterPro; IPR001703; Alphafetoprot.  
CC DR InterPro; IPR000264; Serum albumin.  
CC DR Pfam; PF00273; Serum albumin; 3.  
CC DR PRINTS; PR00802; SERUMALBUMIN.  
CC DR ProDom; PD002486; Serum albumin; 1.  
CC DR SMART; SM00103; ALBUMIN; 3.  
CC DR PROSITE; PS00212; ALBUMIN; 3.  
KW Copper; Lipid-binding; Metal-binding; Repeat; Signal.  
FT SIGNAL 1 18 By similarity.  
FT PROPEP 19 24 By similarity.  
FT CHAIN 25 609 Serum albumin.  
FT DOMAIN 25 206 Albumin 1.  
FT DOMAIN 213 398 Albumin 2.  
FT DOMAIN 405 596 Albumin 3.  
FT METAL 28 28 Copper.  
FT DISULFID 78 87 By similarity.  
FT DISULFID 100 116 By similarity.  
FT DISULFID 115 126 By similarity.  
FT DISULFID 149 194 By similarity.  
FT DISULFID 193 202 By similarity.  
FT DISULFID 225 271 By similarity.  
FT DISULFID 270 278 By similarity.  
FT DISULFID 290 304 By similarity.  
FT DISULFID 303 314 By similarity.  
FT DISULFID 341 386 By similarity.  
FT DISULFID 385 394 By similarity.  
FT

FT DISULFID 417 463 By similarity.  
FT DISULFID 462 473 By similarity.  
FT DISULFID 486 502 By similarity.  
FT DISULFID 501 512 By similarity.  
FT DISULFID 539 584 By similarity.  
FT DISULFID 583 592 By similarity.  
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F675F1A48 CRC64;  
Query Match 63.1%; Score 2387; DB 1; Length 609;  
Best Local Similarity 73.9%; Pred. No. 5e-52;  
Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;  
QY 2 AHKSEVAHRFXDLGSENFKALVLAFAQYLOQCDFEDHVKLVNTEFAKTCVADESSEN 61  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 27 AHKSEIAHRYKDLGKFKYGLVLYTFQYLOKCSYEEHVKLVRVTFASCAKXDESSEN 86  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPRLVPRPVD 121  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 87 CDKSLHTLFGDKLCSLPNFGKYEAMADCCAKQPERNECFLOHKDDNPQLPPFKAEPPD 146  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 122 VMCTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTTECCOAAADKAACLLPK 181  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 147 AMCTAFQENAEAFMGHYLHEVARRRHPYFYGPELLYLDKYTAVLTCCECAADDKGACLTPK 206  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 192 LDELDEGKASAKQRLKASLOKFEGERAFKAWAVARLSORPPKAEFAEVSKLYTDLTKV 241  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 207 LDALKEKALVSAVRQLKCSSMKKFGERAFKAWAVARMSQTFPNADFAEITKLATDLTKV 266  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 242 HTECHGDLLECCADRADLAKYICENODSISSKLKECCCKPPEKSHCIAEVENDEMPAD 301  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 267 TOECCHGDLLECCADRAELAKYICENQASISSKLQCCCKEMLQSKQLAEVEHDDMPAD 326  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 302 LPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVVLRLRLAKTYETTLKCC 361  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 327 LPALTADFVEDKDVCKNYAEAKDVLGTFLEYERRHPDYSVSLRLRLAKTYETTLKCC 386  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 362 AAADPHCYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYFQNALVRYTKKQVSTP 421  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 387 AEADPHACYGHVDFEFPKPLVEEPQNLVKSNCLEKELGEYFQNALVRYTKKQVSTP 446  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 422 TLVEVSRNLGKVGKCKKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTESL 481  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 447 TLVEAARSLGRVGHCCALPEKGRUPCEDYLSALNRLVCLLHEKTPVSEQVTKCCSGL 506  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 482 VNRPCFSALEVDYETVPKFNATFTTFHADICTLSEKERQIKQTALVELVKHKPKATK 541  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 507 VERRECFSALEVDYETVPKFNATFTTFHADICTLSEKERQIKQTALVELVKHKPKQATE 566  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 542 EQLKAVMDPFAFVEKCKKADKCTCFABEGKLVAAASQAAL 583  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 567 EQLKAVMDPFAFVEKCKKADKCTCFABEGKLVAAASQAAL 608  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 17

QEB320  
ID Q6B320 PRELIMINARY; PRT; 583 AA.  
AC Q6B320;  
DT 25-OCT-2004 (TReMBLrel. 28, Created)  
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)  
DE Serum albumin (Fragment).  
GN Name=ESA;  
OS Elephas maximus (Indian elephant).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Proboscidea; Elephantidae; Elephas.  
OX NCBI\_TaxID=9783;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Lazar J., Rasmussen B., Greenwood D.R., Bang I.-S., Prestwich G.D.;  
RT "Elephant Albumin: A Multi-purpose Pheromone Shuttle.";  
RL Chem. Biol. 0:0-0(2004).  
DR EMBL; AY684122; AAT90502.1; -;  
DR GO; GO:0005615; C:extracellular space; IEA.



[6]  
RP SEQUENCE OF 25-44.  
RC TISSUE=Liver;  
RX MEDLINE=93162044; PubMed=1286668;  
RA Giometti C.S., Taylor J., Tollaksen S.L.;  
RT "Mouse liver protein database: a catalog of proteins detected by two-dimensional gel electrophoresis.";  
RL Electrophoresis 13:970-991(1992).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ011413; CAA09617.1; -;  
DR EMBL; AK010025; BAB26650.1; -;  
DR EMBL; BC024643; AAH24643.1; -;  
DR EMBL; BC049971; AAH49971.1; -;  
DR EMBL; M16111; AAA37190.1; -;  
DR EMBL; X13060; CAA31458.1; -;  
DR PIR; A05139; A05139.  
DR HSSP; P02768; 1HK1.  
DR SWISS-2DPAGE; P07724; MOUSE.  
DR MGD; MGI:87991; Alb1.  
DR InterPro; IPR001703; Alphafetoprot.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Serum albumin; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum albumin; 1.  
DR PROSITE; PS00212; ALBUMIN; 3.  
DR Copper; Direct protein sequencing; Lipid-binding; Metal-binding;  
KW Repeat; Signal.  
FT SIGNAL 1 18 By similarity.  
FT PROPEP 19 24  
FT CHAIN 25 608 Serum albumin.  
FT DOMAIN 25 205 Albumin 1.  
FT DOMAIN 212 397 Albumin 2.  
FT DOMAIN 404 595 Albumin 3.  
FT METAL 27 27 Copper.  
FT DISULFID 77 86 By similarity.  
FT DISULFID 99 115 By similarity.  
FT DISULFID 114 125 By similarity.  
FT DISULFID 148 193 By similarity.  
FT DISULFID 192 201 By similarity.  
FT DISULFID 224 270 By similarity.  
FT DISULFID 269 277 By similarity.  
FT DISULFID 289 303 By similarity.  
FT DISULFID 302 313 By similarity.  
FT DISULFID 340 385 By similarity.  
FT DISULFID 384 393 By similarity.  
FT DISULFID 416 462 By similarity.  
FT DISULFID 461 472 By similarity.  
FT DISULFID 485 501 By similarity.  
FT DISULFID 500 511 By similarity.  
FT DISULFID 538 583 By similarity.  
FT DISULFID 582 591 By similarity.  
FT CONFLICT 27 27 H -> D (in Ref. 6).  
FT CONFLICT 33 33 H -> D (in Ref. 6).  
FT CONFLICT 41 41 Q -> I (in Ref. 6).  
SQ SEQUENCE 608 AA; 68692 MW; 292F7CED3A61B4 CRC64;  
62.8%; Score 2378; DB 1; Length 608;

Query Match

Best Local Similarity 72.4%; Pred. No. 8.5e-52;  
Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHREFKOLGEENFKALVLIATAFQAQYLQCPFFEDHVKLVNNEVTEFAKTCVADESAAE 60  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 25 EAHKSEIAHRYNDLGEQHFGLVLIATFQAQYQKCSYDEHAKLVQEVTDFAKTCVADESAA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETGTGEMADCAKQEPERNECFLOHKDNDPNLPRIVREV 120  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCTKQEPERNECFLOHKDNDPNLPPFERPEA 144  
QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFAPELLYFAPELLYAEQYNEIUTQCAEADKESCLTP 180  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 145 EAMCTSFKENPTTFMGLYHVARRRHPYFAPELLYAEQYNEIUTQCAEADKESCLTP 204  
QY 181 KLDELDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRPFAEVSQKLVTDLT 240  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 205 KLDGKVKALVSSVVRQMKCSMQKFGERAFKAWAVARLSQTFPNADFAEITKLTATDLTK 264  
QY 241 VHTCCHGDLLECADDRAADLAKYICENQDSISSKJKECCPKLLEKSHCIAEVDEMPA 300  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 265 VNKECCHGDLLECADDRAELAKYMCENQATISSKJQTCDCPKLLKKAHCLSEVEHDTWPA 324  
QY 301 DLPSLAADFVESKOVCKNYAEAKOVFLGMLFLEYVARRHPDYSVLLLLAKTYETTLK 360  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 325 DLPAIADDFVEDQEVCKNYAEAKOVFLGMLFLEYVARRHPDYSVLLLLAKTYETTLK 384  
QY 361 CAADAPHECYAKVDFEFPKPLVEEPONLIKQNCLEFQELGEYKFNQNALVRYTKVPQVST 420  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 385 CAEANPPACYGTVLAEFQPLVEEPQNLVKTNCDLKYLGEYGFQNALVRYTKQAPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKCHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 445 PTLVEAARNLGRVGTKCTLPEDQRLPCVEDYLSAILNLRVCLLHHEKTPVSEHVTCCSGS 504  
QY 481 LVNRPFCPSALEVDYTPKFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 505 LVNRPFCPSALTVDYTPKFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 564  
QY 541 KEOLKAVMDDFAAVFVEKCKKADDDKTCFAEGCKLVAASQAAL 583  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 565 AEQLKTVMDDFAAFLDTCCKADKDKTCFSTGPNLVTRCKDAL 607  
RESULT 19  
Q8C7H3 PRELIMINARY; PRT; 608 AA.  
ID Q8C7H3  
AC Q8C7H3; PRELIMINARY; PRT; 608 AA.  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Mus musculus adult male liver tumor cDNA, RIKEN full-length enriched library, clone:C730030P03 product:albumin 1, full insert sequence.  
DE Name=Alb1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).

```

[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukushida K., Furuno M., Haneagaki T., Hara A., Hashizume W.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki J., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK050248; BAC34145.1; -.
DR HSSP; P02768; 1HK1.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001703; Alphafetoprot.
DR InterPro; IPR001264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00803; AFETOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

Query Match 62.7%; Score 2374; DB 2; Length 608;
Best Local Similarity 72.2%; Pred. No. 1.1e-51;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVIAFAQYLQCCPFEDHVLKLVNEVTEFAKTCVADESAAE 60
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
RT 25 EAHKSEIAHRYNDLGEQHFGLVIAFQYLQCKSYDESHAKLVQSVTDFAKTCVADESAA 84
RL :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLVRPEV 120
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
RT 85 NCDKSLHTLFGDKLCAIPLNLRNSELADCCCTKQEPERNECFLOHKDNDPNLPPPERPEA 144
RL :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 121 DVMCTAFHDNEETFLKKLYETARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
RL :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki K.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura K.,  
 RA Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh K., Sakai K., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK050644; BAC34360.1; -;  
 DR HSSP; P02768; IHKL.  
 DR MGD; MG1:87991; Alb1.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR InterPro; IPR001703; Alphafetoprot.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 3.  
 DR PRINTS; PR00803; AFETOPROTEIN.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 FT NON\_TER  
 SQ SEQUENCE 576 AA; 65002 MW; F85733B99AE37F04 CRC64;  
 Query Match 61.7%; Score 2336; DB 2; Length 576;  
 Best Local Similarity 72.2%; Pred. No. 8.4e-51;  
 Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;  
 QY 9 HRFKDLGENFKALVLIAPQYLQCFPDHVKLVNEVTEPAKTCVADSAENCKSLHT 68  
 DB 1 NRYNDLGEHFHGLVLIAPQYLQCFPDHVKLVNEVTEPAKTCVADSAENCKSLHT 60  
 QY 69 LFGDKLCTVATRETYGEMADCCAKOEPNERECFLQHKDDNPNLPLVRPEVDVMTAFH 128  
 DB 61 LFGDKLCAIPNRENYGELADCCOEPNERECFLQHKDDNPNLPLVRPEVDVMTAFH 120  
 QY 129 DNEETFKKLYEIAARRHPVYAPPELLFPKRYKAAFTCCQAADKAACLLPKLDELDRDE 188  
 DB 121 ENPTTFMGVHLHEVARRHPVYAPPELLYAEQYNEITLTOCCAEADKESCLTPKLDGVKEK 180  
 QY 189 GKASSAKQRLKASLQKQGERAPKAWARLSQRFPAKPAFVSKVLDTLTKVHTCCCHG 248  
 DB 181 ALVSVVRQRMKSSQKQGERAPKAWARLSQTFPNDAFETIKLATDLTKVNEKCCCHG 240  
 QY 249 DLLECADDRADIATKYICENQDTSKGLKECCEKPLEKSHCIAVENDEMPADLPISLAAD 308  
 DB 241 DLLECADDRAEALAKYCNENQATISSKLTQCCDCKPLLKKAHCUSEVEHDTMPADLPISLAAD 300

QY 309 FVESKDYCKYAEAKDYFLGFLMFLYARRHPDYSVILLRLAKTYETTLKCCCAADPHE 368  
 DB 301 FVEDQECCKYAEAKDYFLGFLMFLYARRHPDYSVILLRLAKTYETTLKCCCAADPHE 360  
 QY 369 CYAKVFDEFKPLVEEPQNLKQNCFLFQOLGEYKFNALLVRYTKVQVSTPTTLVEYSR 428  
 DB 361 CYGVTLAEFQPLVEEPKLVNTCDLYEKLGEYGFQNALVRYTKQAPQVSTPTTLVEAAR 420  
 QY 429 NLGKVGSKCCXHPHPEAKRMPCAEYLSVVLNOLCVLHEKTPVSDRVTKCTCTSLVNRPCF 488  
 DB 421 NLGRVGTCKCTLPEDQRLPCVEDYLSALINRVCLLHEKTPVSEHVTKCCSGSLVRRPCF 480  
 QY 489 SALEVDETYVPKFNAPETTHADICTLSEKEROIKKOTALVELVHKHPKATKEOLKAVM 548  
 DB 481 SALTVDETYVPKFNAPETTHADICTLSEKEROIKKOTALVELVHKHPKATKEOLKAVM 540  
 QY 549 DDFAAVFEKCKCAKDDKETCFAEEGKLVAAQAAL 583  
 DB 541 DDFAQFLDTCCCKAADKDTCFSTEGPNLVTRCKDAL 575  
 RESULT 21  
 Q86YGO  
 ID Q86YGO PRELIMINARY; PRT; 417 AA.  
 AC Q86YGO;  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE ALB protein.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC041789; AAH41789.1; -;  
 DR HSSP; P02768; 1E7B.  
 DR GO; GO:0005615; C:extracellular space; IEA.  
 DR GO; GO:0005386; F:carrier activity; IEA.  
 DR GO; GO:0008810; P:transport; IEA.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; Serum albumin; 2.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 2.  
 DR PROSITE; PS00212; ALBUMIN; 2.



SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;

Query Match 54.1%; Score 2045.8; DB 2; Length 417;  
Best Local Similarity 67.2%; Pred. No. 7.6e-44;  
Matches 393; Conservative 0; Mismatches 0; Indels 192; Gaps 1;

QY 1 DAHKSEVAHFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
DB |||||  
QY 25 DAHKSEVAHFKDLGEEN- 42  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 120  
DB ----- 42  
QY 43 ----- 42  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYVAPPELLFFAKRYKAATFCCQAADKAACLLP 180  
DB ----- 42  
QY 181 KLDELDEGRKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEVSKLVTDLTK 240  
DB -----FKAWAVARLSORFPKAEVSKLVTDLTK 72  
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB ||||| 73 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 132  
QY 301 DLPSLAADFVESKDVCKNRYAEAKDVFGLMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 360  
DB ||||| 133 DLPSLAADFVESKDVCKNRYAEAKDVFGLMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 192  
QY 361 CAADPHECVAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
DB ||||| 193 CAADPHECVAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPOVST 252  
QY 421 PTLVEVSRLNKGKVGSKCKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
DB ||||| 253 PTLVEVSRLNKGKVGSKCKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 312  
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB ||||| 313 LVNRRPCFSALEVDVETVPKFEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 372  
QY 541 KEQLKAVMDDFAAFVEKCCCKADDKTCFAEBGKKLVAASQAALGL 585  
DB ||||| 373 KEQLKAVMDDFAAFVEKCCCKADDKTCFAEBGKKLVAASQAALGL 417

RESULT 22  
Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE ALB protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.B.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madao A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Kzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035969; AAH35969.1; -;  
DR HSSP; P02768; INSU.  
DR GO; GO:0005615; C:extracellular space; IEA.  
DR GO; GO:0005386; F:carrier activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000264; Serum albumin.  
DR Pfam; PF00273; Serum albumin; 2.  
DR PRINTS; PRO0802; SERUMALBUMIN.  
DR SMART; SM00103; ALBUMIN; 2.  
DR PROSITE; PS00212; ALBUMIN; 2.  
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 51.5%; Score 1950.7; DB 2; Length 396;  
Best Local Similarity 63.6%; Pred. No. 1.7e-41;  
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAHKSEVAHFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60  
DB ||||| 25 DAHKSEVAHFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 120  
DB ||||| 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRPEV 144  
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYVAPPELLFFAKRYKAATFCCQAADKAACLLP 180  
DB ||||| 145 DVMCTAFHDNEETFLKKYL----- 163  
QY 181 KLDELDEGRKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEVSKLVTDLTK 240  
DB ----- 164  
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB ----- 164  
QY 301 DLPSLAADFVESKDVCKNRYAEAKDVFGLMFLYEVARRHPDYSVVLLRLAKTYETTTLEK 360  
DB -----YETTTLEK 171  
QY 361 CAADPHECVAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
DB ||||| 172 CAADPHECVAKVDFEKPFLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPOVST 231  
QY 421 PTLVEVSRLNKGKVGSKCKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
DB ||||| 232 PTLVEVSRLNKGKVGSKCKCKHPEAKRMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 291  
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540  
DB ||||| 292 LVNRRPCFSALEVDVETVPKFEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 351  
QY 541 KEQLKAVMDDFAAFVEKCCCKADDKTCFAEBGKKLVAASQAALGL 585  
DB ||||| 352 KEQLKAVMDDFAAFVEKCCCKADDKTCFAEBGKKLVAASQAALGL 396

RESULT 23





## RESULT 24

```
Q8JIA9 PRELIMINARY; PRT; 527 AA.
AC Q8JIA9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Chordata) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RP Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF375971; AAM46104.1; -.
DR HSP; P02768; I57B.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; Alpha-fetoprot.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00803; Serum albumin; 3.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER
SQ SEQUENCE 527 AA; 59711 MW; C62B7998387F5929 CRC64;

Query Match 34.3%; Score 1296.7; DB 2; Length 527;
Best Local Similarity 45.9%; Pred. No. 1.9e-24;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

Qy 62 CDKSLHTLPGDKLCTVATLRETYGEMADCCAKQPERNECFLOHNDPN-LRPLVRPEV 120
Db 5 CLKSLDTIFDEICHEEGFAKY-DLAACCAKAEVERKECLLAHKNAATPGFIPAFQPGI 63

Qy 121 DVMTAFHDNEETPLKYLVEIARRHPYFVAPELLFFAKRYKAFTCCQAADKAACLLP 180
Db 64 EVSKULQDDRLLTLGNYIYEVARHPYLQVPVFATASYDEALTKCCQTADKATCFHP 123

Qy 181 KLDELDEGKASSAKQRLKCSAQKGFGERAFKAWARLSQRFPKABFAEVSKLVTDLTK 240
Db 124 RIPPLIEYKWSNGIQENTCGILKKGERTLKATKLQMSQKPKADFATINKLVEDITH 183

Qy 241 VHTCCGHDLLCADRADLAKYICENQDSISSKLECCBKPILLEKSHCIAEVENDMPA 300
Db 184 MHTCCRGDTLECLRDREALTEYTCSHKDAISSKLPCTCCESVLERGECIVRLNDDKPA 243

Qy 301 DLPSLAADFVESKDVCNVAEAKDVLGMFLYEYARHPDYSVVLLRLAKTVETTLK 360
Db 244 DLSERIAEYEDHPVCDHLAKEQDAFLAKFLYEYSRRHPPELSTQILGVGKGQYQLLERC 303

Qy 361 CAADAPHECAKVPDEFKPLVEEPQNLIKNCELFEOQLGEYKFQONALLVRYTKVPQVST 420
Db 304 KKTNDPPECYQGAEDLKKHIAQFQELVQNCNDLYTLGGYLFHALLIYRKMPQLTS 363

Qy 421 PTLVEVRNIGKVGSKCKGKPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
Db 364 EELIFYTR-ITKAASRCCEVSVDKLPCTEGYDFVLGQICQHRHSSINNVNQCSCNS 422

Qy 481 LVNRRPCFSALEVDTVPKFEFNATETFEADICTLSEKEROIKQOTALVELVKHKPKAT 540
Db 423 YALRSLCITSLGGDEKVPPIEFSADLFTFFEDLCHAAQDKLQERQOMIYNLVKHKENIT 482

Qy 541 KEQLKAYMDDFAAFVEKCKADKCTCFABEGKKLVAASQAAL 583
Db 483 KEQLQTVFGGFTKWKTEKCCRAEDHEACFGEBGPKLVAESQTAL 525
```

## RESULT 25

```
FETA_PANTR STANDARD; PRT; 609 AA.
AC Q28789;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
DE Name=AFP;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=96032345; PubMed=7557431; DOI=10.1016/0378-1119(95)00303-N;
RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczak A.;
RT "the chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to that of gorilla but distinct differences from that of human.";
RL Gene 162:213-220(1995).
CC -I- FUNCTION: Binds copper, nickel, and fatty acids as well as, and bilirubin less well than, serum albumin.
CC -I- SUBUNIT: Dimeric and trimeric forms have been found in addition to the monomeric form (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and yolk sac.
CC -I- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -I- SIMILARITY: Contains 3 albumin domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; U21916; AAA91641.1; -.
DR PIR; JC4258; JC4258.
DR HSP; P02768; I57B.
DR InterPro; IPR001703; Alpha-fetoprot.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Serum albumin; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Copper; Glycoprotein; Metal-binding; Nickel; Repeat; Signal; Sulfation.
FT SIGNAL 1 18 By similarity.
FT CHAIN 19 609 Alpha-fetoprotein.
FT DOMAIN 20 205 Albumin 1.
FT DOMAIN 212 397 Albumin 2.
FT DOMAIN 404 595 Albumin 3.
FT METAL 22 22 Copper and nickel (By similarity).
FT DISULFID 99 114 By similarity.
FT DISULFID 113 124 By similarity.
FT DISULFID 148 193 By similarity.
FT DISULFID 192 201 By similarity.
FT DISULFID 224 270 By similarity.
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
```

FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 251 251 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 609 AA; 68741 MW; C032987CAD0B672B CRC64;  
Query Match 33.2%; Score 1256.7; DB 1; Length 609;  
Best Local Similarity 39.9%; Pred. No. 3.1e-23;  
Matches 237; Conservative 113; Mismatches 231; Indels 13; Gaps 3;  
Oy 4 KSEVAHR-----FKDLGSENFALVLIAPAOYLOQCPEDHVKLVNEVTEPAKT 52  
Db 17 ESRTLRHNEYGIASLDYSQYTAELNLTDLATIPFAQVQVATYKESKMKVDALTAIEK 76  
Oy 53 CVADESANCKSLHTLPGDKLCTVATRETYGEMADCCAKQEPNERCEFLQKDDNP-N 111  
Db 77 PTGDEQSAGCLENQLPAFLLELCREKEILEKYGH-SDCCSQSEGRHNCFLAHRKPTPAS 135  
Oy 112 LPLRVPRVDVMTAFHNEFTLKKLYETARRHPYVAPPELLFAFKRYKAAFTCCCOA 171  
Db 136 IPFQVPEPVTSCAYEDRETTFMKNFYIETARRHPFLYATILLWAARYDKIIPSCCKA 195  
Oy 172 ADKAACLLPKLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEPAEV 231  
Db 196 ENAVECFQTKAATVTKELRESSLLNQHACVMKNFGTRTFQAITVTKLSQKFTKVNFTET 255  
Oy 232 SKLVTDLTVKHTECHGDLLECADRADLAKYICENQDSISKLKCECKELLESKSHCIA 291  
Db 256 OKLVLDVAHVHCHCRGVDLCLDQGEKIMSYICSDQDTLSNKITECKLTLTLEGGQCII 315  
Oy 292 EVENDEMPADLPSLAADFVSKDVCNVAEAKDVLGMLFVYARRHPDYSVLLLRALK 351  
Db 316 HAENDEKPEGLSPNLRFLGDRDFNQFSGSEKNIFLASFVHYSRRHQLQAVSVILRVAK 375  
Oy 352 TVETLEKCCAAADPHCEVAKVDFDKPLVBEPMQNLKQCELPQLGXYFQNALLVRY 411  
Db 376 GYQELLEKCFQTEPLEQDQGEELQYIQESQALAKRSCGLFKLGEYLYQNAFLVAY 435  
Oy 412 TKKVPQVSTPTLAVESRNLKGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 471  
Db 436 TKKAPQLTSSSELMAITRMAATAATCCQLSEDKLLACGGAADIIGHLCIRHETTPVNP 495  
Oy 472 RVTKCTESLVNRRCFSALEVDVTPKFEFNAETFTHADICTLSEKERQIKKQTALVE 531  
Db 496 GVGQCCTSSYANRRFCFSLVVDVTPVPPAFSDDKPEIFHKDLCQAGQVALQTMKQEFFLN 555  
Oy 532 LVKHKPKATKEOLKAVMDDAFAFVCKCKADDKETCFABEGKKLVAASQAALGL 585  
Db 556 LVKQRPQITEQEAEVIAVDFSLLEKCCQGEVCFABEGQKLISKTEAALGV 609  
RESULT 26  
FETA HUMAN  
ID FETA HUMAN STANDARD; PRT; 609 AA.  
AC P02771;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).  
DE Name=AFP;  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=83273664; PubMed=6192439;  
RA Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;  
RT "Primary structures of human alpha-fetoprotein and its mRNA";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608 (1983).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=87185438; PubMed=2436661;  
RA Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;

"Structure, polymorphism, and novel repeated DNA elements revealed by a complete sequence of the human alpha-fetoprotein gene."; Biochemistry 26:1332-1343 (1987).  
[3]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Lung;  
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[4]  
RN SEQUENCE OF 1-28 FROM N.A.  
RX MEDLINE=93278385; PubMed=7684942;  
RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M., Tilghman S., Krumlauf R., Tuddenham E.G.D.;  
RT "A G->A substitution in an HNF I binding site in the human alpha-fetoprotein gene is associated with hereditary persistence of alpha-fetoprotein (HPAFP)"; Hum. Mol. Genet. 2:379-379 (1993).  
[5]  
RN SEQUENCE OF 429-556 FROM N.A.  
RX MEDLINE=83158778; PubMed=6187626; DOI=10.1016/0378-1119(82)90210-4;  
RA Beattie W.G., Dugaiczky A.;  
RT "Structure and evolution of human alpha-fetoprotein deduced from partial sequence of cloned cDNA."; Gene 20:415-422 (1982).  
[6]  
RN PARTIAL SEQUENCE OF 19-609.  
RX MEDLINE=91242409; PubMed=1709810;  
RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F., Ceccarini C., Terrana B.;  
RT "Human alpha-fetoprotein primary structure: a mass spectrometric study."; Biochemistry 30:5061-5066 (1991).  
[7]  
RN PRELIMINARY SEQUENCE OF 19-35.  
RX MEDLINE=77242506; PubMed=70228; DOI=10.1016/0005-2795(77)90198-2;  
RA Yachnin S., Hsu R., Heinrichson R.L., Miller J.B.;  
RT "Studies on human alpha-fetoprotein. Isolation and characterization of monomeric and polymeric forms and amino-terminal sequence analysis."; Biochim. Biophys. Acta 493:418-428 (1977).  
[8]  
RN PRELIMINARY SEQUENCE OF 19-38.  
RX MEDLINE=78001760; PubMed=711198;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Comparative chemical structures of human alpha-fetoproteins from fetal serum and from ascites fluid of a patient with hepatoma."; Cancer Res. 37:3663-3667 (1977).  
[9]  
RN PRELIMINARY SEQUENCE OF 19-39.  
RX MEDLINE=75018719; PubMed=4138095;  
RA Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M., Kontinen A.;  
RT "Alpha fetoprotein: structure and expression in man and inbred mouse strains under normal conditions and liver injury."; Johns Hopkins Med. J. Suppl. 3:249-255 (1974).

[10]  
RN GENE STRUCTURE.  
RP MEDLINE=85182629; PubMed=2580830;  
RA Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,  
RA Tamaoki T.;  
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'  
RT flanking region.";  
RL J. Biol. Chem. 260:5055-5060(1985).  
RN [11]  
RP METAL-BINDING.  
RX MEDLINE=79001617; PubMed=80265;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
RL Cancer Res. 38:3483-3486(1978).  
RN [12]  
RP BILIRUBIN-BINDING.  
RX MEDLINE=80001710; PubMed=89900;  
RA Aoyagi Y., Ikenaka T., Ichida F.;  
RT "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-  
RT binding ability";  
RL Cancer Res. 39:3571-3574(1979).  
RN [13]  
RP SULFATION.  
RX MEDLINE=86042625; PubMed=2414772;  
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
RT "Tyrosine sulfation of proteins from the human hepatoma cell line  
RT HepG2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
CC -!- FUNCTION: Binds copper, nickel, and fatty acids as well as, and  
CC bilirubin less well than, serum albumin. Only a small percentage  
CC (less than 2%) of the human AFP shows estrogen-binding properties.  
CC -!- SUBUNIT: Dimeric and trimeric forms have been found in addition to  
CC the monomeric form.  
CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and  
CC yolk sac.  
CC -!- DEVELOPMENTAL STAGE: Occurs in the plasma of fetuses more than 4  
CC weeks old, reaches the highest levels during the 12th-16th week of  
CC gestation, and drops to trace amounts after birth. The serum level  
CC in adults is usually less than 40 ng/ml. AFP occurs also at high  
CC levels in the plasma and ascitic fluid of adults with hepatoma.  
CC -!- PTM: Independent studies suggest heterogeneity of the amino-  
CC terminal sequence of the mature protein and of the cleavage site  
CC of the signal sequence.  
CC -!- PTM: Sulfated.  
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M10949; AAA51674.1; -;  
CC EMBL; M10950; AAA51675.1; -;  
CC EMBL; V01514; CAA24758.1; -;  
CC EMBL; M16110; AAB58754.1; -;  
CC EMBL; BC027881; AAH27881.1; -;  
CC EMBL; Z19532; CAA79592.1; -;  
CC PIR; A26624; FPHU.  
CC HSSP; P02768; 1E7B.  
CC GlycoSuiteDB; P02771; -;  
CC Sienna-2DPAGE; P02771; -;  
CC Genew; HGNC:317; AFP.  
CC H-InvDB; HIX0004279; -;  
CC MIM; 104150; -;  
CC InterPro; IPR001703; Alphafetoprot.  
CC InterPro; IPR000264; Serum albumin.  
CC Pfam; PF00273; Serum albumin; 3.  
CC PRINTS; PR00802; SERUMALBUMIN.

ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
KW Copper; Direct protein sequencing; Glycoprotein; Metal-binding;  
KW Nickel; Polymorphism; Repeat; Signal; Sulfation.  
FT SIGNAL 1 18  
FT CHAIN 19 609 Alpha-fetoprotein.  
FT DOMAIN 20 205 Albumin 1.  
FT DOMAIN 212 397 Albumin 2.  
FT DOMAIN 404 595 Albumin 3.  
FT METAL 22 22 Copper and nickel.  
FT DISULFID 99 114  
FT DISULFID 113 124  
FT DISULFID 148 193  
FT DISULFID 192 201  
FT DISULFID 224 270  
FT DISULFID 269 277  
FT DISULFID 289 303  
FT DISULFID 302 313  
FT DISULFID 384 393  
FT DISULFID 416 462  
FT DISULFID 461 472  
FT DISULFID 485 501  
FT DISULFID 500 511  
FT DISULFID 538 583  
FT DISULFID 582 591  
FT CARBOHYD 251 251  
FT VARIANT 570 570  
FT SEQUENCE 609 AA; 68677 MW; 4D4E45820E1C2D4F CRC64;  
FT /FTID=VAR\_012049.  
FT /FTID=CAR\_000070.  
FT N-linked (GlcNAc...).  
Query Match 33.1%; Score 1252.7; DB 1; Length 609;  
Best Local Similarity 39.7%; Pred. No. 3.9e-23;  
Matches 236; Conservative 114; Mismatches 231; Indels 13; Gaps 3;  
Qy 4 KSEVAHR-----FKDLGEENFKALVLIATAFYQQLOQCPEDHVKLVNEVTEFAKT 52  
Db 17 ESRTLHNEYGIASILDYSQCTAEISLADLATIFFAQVQEAITYKESKMKVDALTAIEK 76  
Qy 53 CVADESSENCKSLHTLFGKLCIVATLRETYGEMADCCAKQEPERNECFLOHDDNP-N 111  
Db 77 PTGDEQSGGLENLQPAFLPEELCHEKEILEKYGH-SDCCSQSEGRHNCFLAHKKPTPAS 135  
Qy 112 LRLVLRPEVDVMTAFHDNEETPLKYLIELTARRHPYFYAPPELLFFAKRYKAAATECCOA 171  
Db 136 IPIFOVPEPTVSCAYEEDRETFWNKFIYELIARRHPFLYAPTILLWAARDKIIPSCCKA 195  
Qy 172 ADKAAICLLPKLDELDRDEGKASSAKQRLKCSAQKQGERAFKAMAVARLSQRPKAEFAV 231  
Db 196 ENAVECFQTKAAATVTKELRESSLLNQHACAVMKNGFTRTFQAITVKLSQKFTKVNFTET 255  
Qy 232 SKLVTDLTQVHTECHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIA 291  
Db 256 QKLVLDVAHVHEHCRGDVLDCLQDGEKIMSYICSQDDTLNSNKITECCKLTTLERGGCII 315  
Qy 292 EVENDENPADLPSLAADPFVESKKNVABAKVFLGMLFYEVARRPDVSVLLLRLLAK 351  
Db 316 HAENDERPEGLSPNINRPLGDRDNQFSSEKKNIFLASFVHEYRRRPLQAVSVILRVAK 375  
Qy 352 TYETTLKCCAAADPHECYAKVDFDEKPLVVEEPQNLIKONCELFQELGEYKFQNALVRY 411  
Db 376 GYQELLEKCFQTNPLEQCQDQGBEELQYIQESQALAKRSQGLQKLGYYLQNAFLVAY 435  
Qy 412 TKKVPQVSTPTLVSVSNLKGKCKHPKAKMPCAEYLSVVLNQLCVLHEKTPVSD 471  
Db 436 TKKAPOLTSSELMAITRKMAATAATCCQLSEDKLLACGEGAADIIGHLICIRHEMTVPNP 495  
Qy 472 RVTKCCTESLVNRRPCFSALEVDVETYPKPFNAETFTFHADICTLSEKERQIKQTALVE 531  
Db 496 GVGQCCTSSVANRRPCFSSLVVDVETIYPPAFSDDKFIHFKDLCCQAQGVALQTMKEFLIN 555  
Qy 532 LVKHKPKATKEQLKAVMDDFAAAFVEKCKCKADDKETCFAEKGLKLVAAASAALGL 585

**RESULT 28**

FETA GORGO STANDARD; PRT; 609 AA.

ID FETA\_GORG  
AC P28050;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 05-JUN-2004 (Rel. 44, Last annotation update)

DQ Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoproteine).

DE Name=APP;  
GN Gorilla gorilla gorilla (Lowland gorilla).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
OX NCBI\_TaxID=9595;

[1]  
SEQUENCE FROM N.A.  
MEDLINE=91169517; PubMed=1706310;  
RYAN S.C., Zielinski R., Dugaiczky A.;  
RT "Structure of the gorilla alpha-fetoprotein gene and the divergence of primates";  
RL Genomics 9:60-72(1991).

-! FUNCTION: Binds copper, nickel, and fatty acids as well as, and bilirubin less well than, serum albumin.  
CC -! SUBUNIT: Dimeric and trimeric forms have been found in addition to the monomeric form (By similarity).  
CC -! SUBCELLULAR LOCATION: Secreted.  
CC -! TISSUE SPECIFICITY: Plasma.  
CC -! PTM: Sulfated (By similarity).  
CC -! SIMILARITY: Belongs to the ALB/APF/VDB family.  
CC -! SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; M38272; AAA73520.1; --  
PIR; A37970; FPGO.  
DR HSFP; P02768; IHK2.  
DR InterPro; IPRO01703; Alphafetoprot.  
DR InterPro; IPRO00264; Serum\_albumin.  
DR Pfam; PF00273; Serum\_albumin; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN\_3.  
DR PROSITE; PS00212; ALBUMIN; 2.  
KW Copper; Glycoprotein; Metal-binding; Nickel; Repeat; Signal; Sulfation.

FT SIGNAL 1 18 By similarity.  
. FT CHAIN 19 609 Alpha-fetoprotein.  
FT DOMAIN 20 205 Albumin 1.  
FT DOMAIN 212 397 Albumin 2.  
FT DOMAIN 404 595 Albumin 3.  
FT METAL 22 22 Copper and nickel (By similarity).  
FT DISULFD 99 114 By similarity.  
FT DISULFD 113 124 By similarity.  
FT DISULFD 148 193 By similarity.  
FT DISULFD 192 201 By similarity.  
FT DISULFD 224 270 By similarity.

```
FT DISULFID 269 277 By similarity.
FT DISULFID 289 303 By similarity.
FT DISULFID 302 313 By similarity.
FT DISULFID 384 393 By similarity.
FT DISULFID 416 462 By similarity.
FT DISULFID 461 472 By similarity.
FT DISULFID 485 501 By similarity.
FT DISULFID 500 511 By similarity.
FT DISULFID 538 583 By similarity.
FT DISULFID 582 591 By similarity.
FT CARBOHYD 251 251 N-linked (GlcNAc...) (potential).
SQ SEQUENCE 609 AA; 68697 MW; EBAE548377DB60B8 CRC64;

Query Match 32.9%; Score 1245.7; DB 1; Length 609;
Best Local Similarity 39.4%; Pred. No. 5.8e-23;
Matches 234; Conservative 115; Mismatches 232; Indels 13; Gaps 3;

Qy 4 KSEVAHR-----FKDLGENFKALVLIAPFAVLOQCPPEHDHVKLVNEVTEFAKT 52
Db 17 ESRTLHRENYGIASILDSYQCTAIEISLADLATIFFAQFVQEAITYKSVKVKDALTAIEK 76
Qy 53 CVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-N 111
Db 77 PTGDEQAGGLENLQAPAFLELCHKEILEKYG-LSDCSSQSEGRHNCFLAHKKPTPAS 135
Qy 112 LPLRVREVDVMTAFHDNEETFLKKYLITAREHPHYFAPELLFFAKRYKAAFTCCQA 171
Db 136 IPLFQVPEPTVSCAYBEDRETFTNKFYIETAREHPFLYATILLWAARVDKIIPSCCKA 195
Qy 172 ADKAACLLPKLDELDEGKASSAKQRLKCSAQKQGERAFKAMAVARLSORFPKAEFAEV 231
Db 196 ENAVECQTAAATVKELRESSLLNQHACVMKNFGTRTQAIIVTKLSQKFVKNVTEI 255
Qy 232 SKLVTDITKVHTECHGDLLECADDRADLAKYICENQDSISSKLECEKPLEKSHCIA 291
Db 256 QKLVLDVAHVEHCRCGDLDCLODGEKINSYICSQDITLSNKITECKLTTLERGQCII 315
Qy 292 EVENDEPADLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVVLLRLAK 351
Db 316 HAENDEKPEGLSPNLNRDGRDNFQFSSGKKNIFLASFVHEYSRRHPQLAVSVILRAK 375
Qy 352 TYETTLKCCAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQELGEYKFNALLVRY 411
Db 376 GYQELLECKQTENPLSCQDKGEELQYIQESQALAKRSCGLFQKLGEYLLQNAFLVAY 435
Qy 412 TKKVPQVSTPTLVEVSRLNLGKSGCKHPKAPKMPCAEDYLSVLNLQCLVLHEKTPVSD 471
Db 436 TKKAPQLTSSLMATRKMAATAATCCQLSBDKLLACGEGAADIIGHLCIRHEMTVPNP 495
Qy 472 RVTCKCTESLVNRRPCFSALEVDVYVPKFNFAETFTFHADICTLSEKERQIKQTALVE 531
Db 496 GVGQCCTSSYANRRPCFSSSLVVDVYPPAFSDDKFIHFKDLCAQGVALQTMKQEFLLN 555
Qy 532 LVKHKPKATKQKLVAMDDFAAFVEKCKADKDETCFAEEGKLVAAASQAALGL 585
Db 556 LVKQKQITEQLTEVIADFSGLLEKCCQGEQVEVCAEEGQKLSIKTRALGV 609
```

RESULT 29

```
Q8MJ76
ID Q8MJ76 PRELIMINARY; PRT; 609 AA.
AC Q8MJ76
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-fetoprotein.
GN Name=afp;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Furuichi M., Neo S., Hisasue M., Teuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Yamada T.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -.
DR HSP; P02768; IHK2.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001703; Alpha-fetoprot.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 3.
DR PRINTS; PR00803; AFEPTOPROTEIN.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B8250C5AF2AF0 CRC64;

Query Match 32.8%; Score 1242.8; DB 2; Length 609;
Best Local Similarity 40.5%; Pred. No. 6.9e-23;
Matches 231; Conservative 113; Mismatches 225; Indels 2; Gaps 2;

Qy 16 EENFKALVLIAPFAVLOQCPPEHDHVKLVNEVTEFAKTCTVADESAENCDKSLHTLFGDKLC 75
Db 40 EMNLVDLATIFFAQFVQEAITYKSVKVKDILTVIEKSTGSEQPGGCLNLQAPAFLEIEIC 99
Qy 76 TVATLRETYGEMADCCAKQEPERNECFLOHKD-DNPNLPLRVREVDVMTAFHDNEETF 134
Db 100 HEKEISEKYG-LADCCSQSEERHNCFLAHKKAAPPISPPFVABPVTSKAYEENRDMF 158
Qy 135 LKLYVBIARRHPHYFAPELLFFAKRYKAAFTCCQAADKAACLLPKLDELDEGKASSA 194
Db 159 MNRVYIARIARHPFLVAPTILSLAAHYGKILPICCKAENAVECFQYTSILITELRESSL 218
Qy 195 KQRLKCSAQKQGERAFKAMAVARLSORFPKAEFAVSKLVTDITKVHTECHGDLLECA 254
Db 219 LNQHICAVMRNFGPRTFRAITVTKLKQSKFSAKNTFETQKLVLDVAHTEECRCGNVLECL 278
Qy 255 DRRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDENPADLPSLAADFVESKD 314
Db 279 QDGEKINSYICSQDITLSNKIADCCCKLPLELQGCIIHAENDGKPEGLSPNLNRFLBERD 338
Qy 315 VCKNYAEAKDVLGMLFYEARHPDYSVVLLRLAKYETTLKCCAAADPHECYAKVF 374
Db 339 FNGFSREKDLFWARFYEYSRRHTKLVAVPVLVAKGYQELLEKCSQSENFLECDQKE 398
Qy 375 DEFKPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVPQVSTPTLVEVSRLNLGKVG 434
Db 399 EELEKYIQESQALAKRSCGLFQKLGEYLLQNAFLVAYTKKAPQLTPPELMAFTRKMTAA 458
Qy 435 SKCKHPKAPKMPCAEDYLSVLNLQCLVLHEKTPVSDRVTCKCTESLVNRRPCFSALEVD 494
Db 459 ATCCQLSBDKLLACGEGAADIIGHLCIRHEETPINFGVQCCTSSYANRRPCFSSSLVWD 518
Qy 495 ETVPKFNFAETFTFHADICTLSEKERQIKQTALVELVHKPKATKQKLVAMDDFAAF 554
Db 519 ETVPSPFSADKFIHFKDLCAQGVALQTMKQFLINLVKQKQITEQLTELEAVIADFSGL 578
Qy 555 VEKCKKADDETCFAEEGKLVAAASQAALGL 585
Db 579 LEKCCQGEQVEVCAEEGKLVAAASQAALGL 609

RESULT 30
Q8MJ76
ID Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



**THIS PAGE BLANK (USPTO)**